ERIC Document Resume	See Instruction Sheet	DATE OF RESUME	DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER: ED001133
2. ORIGINATOR CONTROL NO(S).: N.A.
3. IDENTIFICATION NO.: CR SAE 8961

4. AUTHOR(S): Rothney, John W. M.

5. TITLE OF DOCUMENT ABSTRACTED:

Studies in the Discovery and Guidance of Superior Students

6. PUBLICATION TITLE: N.A.

7. EDITOR(S): N.A.

8. PLACE OF ORIGIN
ORG. Univ. of Wisconsin
ST. Madison, Wisconsin
CITY, STATE
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE
Cooperative Research Program, Office of Edu., Dept. of HEW, Washington, D. C.

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	63	106

DO NOT WRITE IN SPACE BELO

MICROFILM COPY AVAILABLE
YES ☒ NO ☐

COPYRIGHT GRANTED
YES ☐ NO ☒

COPYRIGHT RELEASE FILED
YES ☐ NO ☐

TYPE OF REPORT
PROPOSAL ☐
IN-PROGRESS ☐
FINAL ☐
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR: experimental study

16. INDEXING TERMS (6-12)

secondary schools	guidance counselors			
student writings	children, gifted	research opportunities		
student attitudes	testing techniques	creativity		

17. ABSTRACT (MAX. 250 WORDS)

A research and guidance laboratory was set up at the University of Wisconsin. The subjects were superior students chosen from over a hundred Wisconsin high schools. The students were tested and counseled at least once a year from ninth grade through graduation.. The laboratory permitted exhaustive studies on the characteristics, development, and guidance of superior students. An initial study pursued the analysis of personal documents written by superior students seemed to provide a way to see, in part, the inner half of an adolescent that does not always become apparent to those who observe and test him.

A second study used oral problem-solving performances of superior students as one means of evaluating the processes of reasoning and providing effective counseling techniques in problem-solving.

A third study showed the correlation between guidance counselors' prediction of post-high school work through personal evaluations and aptitude tests and the students' actual performance in college. The test is as yet incomplete.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION

20. DIVISION OR DEPT.

21. CONTRACT NO.

PROJECT NO.

22. TOTAL COST OF PROJECT

23. DURATION OR PROJECT

U.S. Department of Health, Education, and Welfare
Office of Education

ERIC F — 1

ED001133

932

Cooperative Research Project No. SAE 8961
Studies in the Discovery and Guidance of Superior Students

John W. M. Rothney
University of Wisconsin 1963

REPRINTED BY PERMISSION
U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN PRINTED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

STUDIES IN THE DISCOVERY AND GUIDANCE OF SUPERIOR STUDENTS

Cooperative Research Project No. SAE 8961

John W. M. Rothney

University of Wisconsin
Madison, Wisconsin
1963

The research reported herein was performed pursuant to a contract with the Cooperative Research Program of the United States Office of Education, Department of Health, Education, and Welfare.

PREFACE

This report covers one phase of a long-term project at the Research and Guidance Laboratory of the University of Wisconsin that has required and will require the cooperative efforts of hundreds of students, their parents, their teachers, and members of the faculty of the University of Wisconsin. The Cooperative Research Program has made it possible to carry out certain phases of the program by supplementing the support provided by the University and 91 cooperating schools.

The persons who have contributed to the project are too numerous to list. We are particularly indebted to the 1400 students who have given their time freely while they were students in high school and after they had graduated. The graduate students who acted as research assistants have worked far beyond the call of duty. Professors on the University staff have contributed many hours without cost to the project. Members of the Laboratory advisory committee, particularly Dean Kurt F. Wendt of the College of Engineering and Dean Lindley J. Stiles of the School of Education, have arranged for excellent facilities in which to conduct the studies that have been done at the Laboratory. The Numerical Analysis Laboratory at the University has provided excellent services.

The studies reported in the following pages extended over a three-year span. This is a relatively short period of time for the study of the complex problems one meets in the discovery and guidance of superior students. The data collected in the studies reported here will be used many times as these subjects are followed through into their post-high school educational and vocational experiences. As the "moving pictures" of these subjects develop the "snapshots" obtained during the three-year period under review in the current studies may prove to be more significant than they now appear.

CONTENTS

Page

PREFACE i

CHAPTER I. Background for the Studies Reported in this Volume 1

CHAPTER II. The Use of Personal Documents in Analyses of Attitudes of Superior Students 6

CHAPTER III. Study of Oral Problem-Solving Performances of Superior High School Students 58

CHAPTER IV. Problems in Actuarial and Clinical Prediction of Performances of Superior Students 89

TABLES

CHAPTER II.

1. Analysis of Spelling and Sentence Faults Recognized by Sixty Superior Students in the Language Usage Test of the Differential Aptitude Test Battery 7
2. Comparison of Superior Students' Proficiency in Recognizing Errors of Language Usage with Their Average Error of Written Usage 9
3. Summary of 2800 Categorizations Obtained from One Hundred Tenth Grade Essays 13
4. Percent of Ninth and Tenth Grade Superior Students Whose Essays were Judged to Contain References to Categories to the Extent Indicated 14
5. Rank Order Percentages of Ninth and Tenth Grade Subjects Whose Essays were Judged to Contain References to Categories Indicated 15
6. Percent of Agreement Between Three Judges on Categorization of Responses in Twenty-five Essays 27
7. Extent of Personal Involvement Evident in Essays on "When Teenagers Should be Treated as Adults" 28
8. Extent to which Outline was Followed in Essays on "When Teenagers Should be Treated as Adults" 28
9. Acceptance or Rejection of Criteria in Essays on "When Teenagers Should be Treated as Adults" 29
10. Originality Evident in Essays on "When Teenagers Should be Treated as Adults" 29
11. Degree of Concern Evident in Essays on "When Teenagers Should be Treated as Adults" 30

Page
12. Attitudes Toward Adults Evident in Essays on "When Teenagers Should be Treated as Adults" 31
13. Attitudes Toward Teenagers Evident in Essays on "When Teenagers Should be Treated as Adults" 31
14. Attitudes Toward the Future Evident in Essays on "When Teenagers Should be Treated as Adults" 32
15. Scope of View of Subjects Expressed in the Senior Essay 44
16. Brief Interpretations of Students' Description of the Condition of the World in the Future 45
17. Definiteness of Plans of Superior Students in Their Senior Year of High School 47
18. Concept of Vocational Roles Reported by Senior High School Superior Students 49

CHAPTER III.
1. Percentages of Two Hundred Students in Grades Nine through Twelve who Answered the Exercises Correctly 64
2. Percentages of Fifty Students in each of the Grades Nine through Twelve who Answered the Exercises Correctly . . . 70
3. Percentages and Differences Between Percentages of One Hundred Boys and One Hundred Girls who Answered the Twelve Exercises Correctly 71
4. Mean Times in Seconds and Differences Between the Means of Ninth and Tenth Grade Boys and Girls who Answered the Exercises Correctly 72
5. Numbers, Percentages and Differences Between Percentages of Boys and Girls who Gave a Right or Wrong Answer to the Problems and who Indicated They were Certain of the Correctness of Their Answers 73
6. Numbers, Percentages and Differences Between Percentages of Boys and Girls who Gave a Right or Wrong Answer to the Problems and who Indicated They were Doubtful of the Correctness of Their Answers 74
7. Percentages of Two Hundred Subjects who Responded in the Manner Indicated to the Question, "Did You ever get a Mental Picture of the Problem Situation?" 75
8. Percentages of Two Hundred Subjects who Responded in the Manner Indicated to the Question, "Do You have any Comments to Make About an Exercise of this Type?" 76

CHAPTER IV.
1. Coefficients of Correlation Between Test Scores and High School and College Grade-Point Averages of Eighty-nine Superior Students who Entered Colleges in September, 1961 95
2. Relationship Between Counselors' Predictions of College Grade-Point Averages and Actual Performances of Eighty-nine Superior Students who Graduated from High School in 1961 96

Page

3. Coefficients of Correlation Between Test Scores and High School and First Semester College Grade-Point Averages of Sixty-six Students who Entered the University of Wisconsin in September, 1963 97
4. Relationship Between Counselors' Prediction of First Semester Grade-Point Averages and Actual Performances of Sixty-six Students who Entered the University of Wisconsin in September, 1962 98

CHAPTER I

BACKGROUND FOR THE STUDIES REPORTED IN THIS VOLUME

Because the reports of studies in this volume grew out of a simultaneous attack on several significant problems related to the discovery and guidance of superior secondary school students, it seems desirable to precede the reports with a brief description of the setting in which they were attacked.

The Laboratory, which began as a small cooperative pioneering project of the School of Education and the College of Engineering in February, 1957, is now an officially recognized part of the University of Wisconsin. A committee composed of the Deans of Letters and Science, Engineering, Education, Law, Medicine, Agriculture, and representatives of the School of Commerce and the Extension Division advises the director of the Laboratory. The work on the Madison campus is done by the writer and a staff of twelve graduate students in guidance. At the Wausau branch, which was opened in September, 1960, the work is done by a former staff member of the Madison Laboratory and one graduate student.

The ninety-one cooperating high schools were those which responded to a brief note in the bulletin of the State Department of Public Instruction to the effect that the University would cooperate with schools in experimentation on the discovery and guidance of superior students. The current participating schools provide an excellent sample of the secondary schools of the State in terms of size and geographic distribution. In size, the range is from schools with ten teachers to one with more than a hundred. This distribution is significant in view of the fact that most previous studies of superior students have been carried out in large cities or in large schools, and the bright student in the small school has received scant attention. Students are brought to the Madison Laboratory from some communities that are more than 250 miles distant. The schools provide for all travel expenses of the students and staff members who accompany them.

At the time the Laboratory was opened, a crucial decision about the method of selecting students was required. It would have been possible to direct the schools to send only those ninth grade students whose test scores were above some arbitrarily selected figure. Instead of using methods of selection which put too much emphasis on test scores, the Laboratory staff encourages the schools to employ their own selection procedures. The following suggestions are given to the schools in written form, and they are discussed frequently in conferences and in school faculty meetings.

> Since superior students exhibit their superiority in many ways, no single score can be used as the sole criterion on which they should be selected. Teachers' judgments, evidence of a high level of performance in any of the academic fields, students' written performances, high grades, and evidence of high motivation coupled with high test scores can all be used in the identification of potentially superior students.
>
> In some schools, principals have asked teachers to nominate their superior students. The principal then joins a committee of teachers to review the nominations, considers such other factors as those noted above, and makes a list of all the top students nominated. Frequently the list of most highly rated nominees, along with the reasons for their selection, are returned to the teachers for further consideration.
>
> Combinations of some of the following characteristics are usually apparent in superior students.
>
> 1. Large vocabularies which are used easily and accurately.
> 2. Academic work one or more years in advance of their classes.
> 3. They ask many penetrating questions. They want to know the causes and reasons for things.
> 4. A wide range of interests but they may concentrate heavily on one.
> 5. Express themselves well in writing and speaking.
> 6. Spend much time on special projects of their own such as constructing, collecting and writing.
> 7. Tendency to figure out what is wrong with an activity and show how it can be done better.
> 8. Tendency to give refreshing twists even to old ideas.
> 9. They like to get answers to problems, puzzles, and trick questions.
> 10. Other students tend to turn to them for ideas and suggestions when something must be decided.
> 11. They read a good deal and find satisfaction in thinking about and discussing what has been read.

Selection procedures used by the schools have become more effective over the past four years. Average scores of ninth grade students sent to the Laboratory are at the 95th percentile on Form 1A of the School and College Ability Test. Their general performances in school, on other tests, and on written documents suggest that the schools are selecting students of high potential. Apparently teachers can recognize superior students when they are given assistance in doing so, and statements that commonly appear in the literature to the opposite effect seem to have been premature and based on inadequate data. There is evidence to suggest that the process employed at the Laboratory results in greater awareness by school personnel of the characteristics of the total school populations from which the Laboratory subjects are selected.

The Laboratory can take only a sample of the students in any school, demonstrate procedures with them, and encourage schools to use and expand upon the methods employed. The size of the sample from each school depends on the population of ninth graders, but some flexibility is permitted. If schools, after looking over their students, request an increase in the size of their quotas and even in some cases reduction of them, they are permitted to do so.

The students come to the Laboratory for the first time when they are in the ninth grade. They continue to come for a full day at least once each year until they graduate from high school. During their visits the students are tested and counseled. They write personal documents so that counselors can assess their writing skills and learn more about their background, problems, and plans, and they report on their activities orally and in writing. The total performance of the day is interpreted to the student, and all information obtained is shared with him. They are permitted to attend university classes in the subject area of their choice, and they are permitted to visit various studios and laboratories. When a student is beginning to express interest in a particular subject-field, an interview is arranged for him with an authority in that field, and the implications are discussed later with his counselor.

In some cases the students bring samples of their writing, art work, or other productions to the Laboratory for assessment by experts in their fields of interest. The professors give generously of their time for appraisal of the students' performances and provide constructive criticism and encouragement. At least one poet and several artists and writers of considerable promise have been discovered. All members of the university faculty who have been asked to confer with students, permit visits to their classes, and appraise their work have responded with enthusiasm.

After the student returns to his school, a report on each student is sent to the principal. It describes in two or three pages the students' performances, problems, and plans, and it ends with a number of suggestions which the school is asked to consider. The schools usually require each teacher who has the student in her classes to read the reports and follow the reading with participation in a group session in which they are discussed. In the smaller schools all members of the faculty read the report and participate in the discussions.

Members of the Laboratory staff visit the school for a full day shortly after the reports have been sent to the schools. At this time they meet with parents of all freshmen and seniors. Conferences with parents of sophomores and juniors are held if they are requested. In most cases each father and mother has been seen separately for a short period to permit collection of research data, and then both parents have been brought together for a discussion of their child's performances, problems, and plans. The members of the Laboratory staff also confer with the students' teachers individually or in small groups, and the day is completed with a faculty meeting at which each staff member presents a report on the student with whom he worked. Vigorous discussion of particular students and of general methods of identifying and guiding superior students usually follows.

The Laboratory staff plans to follow these students for many years through their post-high school education and into employment. The first small group of trial subjects who had come to the Laboratory during their four years of high school were graduated in 1960. Follow-up procedures were tried out with them, and control groups of subjects from their college classes were selected. It is anticipated that some 250 new subjects will be added and that approximately the same number will be graduated each year. Continuation of Laboratory practices with those currently in high school and annual follow-up studies of the graduates of each year assures a large population for research purposes.

The Laboratory was designed primarily for research in the discovery and guidance of superior students. It serves also as a place where graduate students in guidance can get supervised practice in counseling and research in guidance. The wealth of longitudinal data permits and will continue to permit exhaustive studies on the characteristics, development, and guidance of the students. It also permits intensive studies of their family and school circumstances.

It was in the setting described above that the three studies reported in the following pages were carried out. It may be observed that the circumstances permit the follow-up of the subjects of these studies for many years and that there will be a continuing study of relationships among performances described in the following pages and the post-high school social, educational and vocational achievements of these subjects.

CHAPTER II

THE USE OF PERSONAL DOCUMENTS IN ANALYSES OF ATTITUDES OF SUPERIOR STUDENTS

It was noted in Chapter I that the students came to the Laboratory once each year while they were in the ninth, tenth, eleventh and twelfth grades. On each visit they write an essay on an assigned topic. Those assigned by year are described below.

Ninth Grade Essay

You will be given this hour to write an essay about yourself entitled, "The Dominant Forces That Have Directed My Life." Perhaps it would help if you keep the following questions in mind while you write.

1. In what ways have certain persons influenced my life most?
2. What other factors have caused me to be the person I am?
3. How will the above two factors aid or hinder me in attaining what I hope to become?

Tenth Grade Essay

The purpose of this assignment is to make you aware of the variety of things that influence us in our "growing up" and in developing plans for the future. Besides being an assignment in writing about a very interesting person, YOU, the autobiography will help you decide "What kind of person am I?" "How did I get that way?" and "What do I hope to become?" Keep these three questions in mind and write freely about yourself. Include anything that you feel helped make you what you are. You will have about an hour to write.

Eleventh Grade Essay

The eleventh grade essay required the students to write on the topic of when teen-agers reach maturity. It was similar to the writing section of the Sequential Tests of Educational Progress. Students were given a series of questions about maturity to guide their writing and were asked to justify their answers.

Twelfth Grade Essay

You will be given one hour to write on the following topic: THE FUTURE AS I SEE IT AND MY PLACE IN THAT FUTURE.

Analyses of the essays served 2 different purposes. In the first place they provided samples of the quality of the students' writing with respect to such matters as spelling and language usage which could be used in counseling. In the second place they provided some basic data for analyses to determine the value of personal documents in the discovery of adjustments and attitudes.

As an illustration of the use of essays as evidence of quality of writing performances the following study was conducted.

Sixty high school freshmen, 30 boys and 30 girls, were chosen alphabetically from 125 superior students who came to the Laboratory during one year. The students were representative of schools varying in size from 100 to 2500 students from all sections of Wisconsin. Data were obtained from both the Differential Aptitude Language Usage Test (Form A) and the essay entitled "The Dominant Forces That Have Directed My Life." In categorizing the superior students' recognition of language usage errors, the ninety-five errors present in the sentence section of the test were placed into the 7 classifications which appear in Table 1. The bases of the classifications were drawn principally from grammar classics and 10 of the basic grammar and composition tests which are used in the eighth and ninth grades in Wisconsin high schools. The "spelling"

Table 1. Analysis of Spelling and Sentence Faults Recognized by Sixty Superior Students in the Language Usage Test of the Differential Aptitude Test Battery

Types of Error	Possible Errors	Average Number of Incorrect Items Recognized Boys	Girls	Total
Spelling	100.0	84.5	88.6	86.6
Verb usage	18.0	12.4	12.3	12.4
Agreement faults	14.0	6.0	7.2	6.6
Pronoun usage	18.0	11.5	12.1	11.8
Adjective-adverb usage	16.0	9.8	10.9	10.8
Capitalization and punctuation	9.0	5.1	5.9	5.5
Violations of acceptable English	17.0	8.2	8.4	8.3
TOTALS		53.0	56.8	63.4

section of the Language Usage Test was analyzed according to the correct and incorrect answers given by the students.

The classification used in the analysis of the test results was also employed in an examination of the essays written by the same sixty students, so that comparative data could be obtained. Since there is a recognized subjectivity in marking essays, just as there is in the construction of standardized tests, the essays were read by the investigator three times each in random order, with a period of five days elapsing between each reading. Fifteen of the essays were also selected randomly and read by a Wisconsin high school teacher experienced in teaching superior ninth grade students. The errors and comments recorded by each reader independently corresponded very favorably. Comparative figures were compiled from the data obtained from both the standardized test results and the essays under the classifications indicated in Table 2.

Computations from the figures in Table 1 indicate that the freshman superior students in this study recognized sixty-three percent of the language usage errors on the standardized test. This percentage compares favorably with the percentage achieved by students who score near the 99th percentile in the Language Usage Test. The superior students found greatest difficulty in recognizing errors concerned with "agreement faults" and "violations of acceptable English." In both areas just less than half of the items were correctly recognized.

The nature of the essay data understandably prevented the establishment of a ceiling of error for each of the classifications. Examination of the comparative figures in Table 2 seem to indicate, however, that superior students wrote with less incidence of error than could have been predicted by the difficulty they had experienced in the identification of language usage errors. The most important figures cited in Table 2 relate to "prounoun usage" and "violation faults." There was a marked difference between the types of items used by the test authors in these areas and the actual written errors of these types found in the essays. The superior students did not seem so bothered with the test authors' emphasis on the correct form of pronouns as they were with the accurate reference in their usage. It is interesting that the ten superior students who achieved the best scores on the standardized test items concerned with "violation faults" made more mistakes in this same area than the entire group from which they were drawn. Conversely, the ten superior students who erred most often in the recognition of "violation faults" wrote with less difficulty in this area than the entire group. The essays also showed that although the superior students made fewer types of error under each classification, their errors were repeated frequently.

Table 2. Comparison of Superior Students' Proficiency in Recognizing Errors of Language Usage with Their Average Error of Written Usage

Types of Error	No. of Possible Errors	Thirty Boys	Thirty Girls	Top Ten Students*	Bottom Ten Students*	All Students
Verb usage						
Error of recognition (test)	18	5.5	5.5	3.6	7.5	5.5
Error of usage (essay)		0.9	0.4	0.2	0.9	0.7
Agreement faults						
Error of recognition (test)	14	7.9	6.9	4.7	9.8	7.4
Error of usage (essay)		0.5	0.8	0.4	0.9	0.7
Pronoun usage						
Error of recognition (test)	18	6.3	5.9	2.3	11.1	5.1
Error of usage (essay)		1.8	2.1	1.7	2.6	1.9
Adjective-adverb usage						
Error of recognition (test)	16	6.1	5.1	1.4	9.3	5.6
Error of usage (essay)		0.7	0.6	0.5	0.8	0.7
"Violation" faults						
Error of recognition (test)	17	8.8	8.7	4.7	11.4	8.7
Error of usage (essay)		4.5	5.1	6.0	4.3	4.8

*The terms "top ten" and "bottom ten" are based upon the ten highest and the ten lowest scores made by students participating in this study.

Some aspects of the compositions did not permit statistical comparisons, yet they revealed enlightening facts about the writing problems of superior students. Some found difficulty in organizing their work and the essays featured unsatisfactory paragraph planning. Sentences were occasionally unrelated to the central theme, and showed a lack of unity, coherence, and clarity. The chief sentence faults were incomplete and rambling structures. It should be noted, nevertheless, that the above comments refer more to inadequacies in meaningful expression than in technical grammatical skills since fifteen of the students wrote technically adequate compositions which were uninteresting in content. An overall appraisal of the entire population revealed a greater technical than expressive proficiency, although both needed improvement.

From the data presented, it would seem that the original hypotheses have been partially substantiated. Evidence of language usage stems from performance with words and not from recognition of items of error. The test data derived from the D.A.T. Language Usage Test did not provide meaningful scores that were predictive of superior students' performances in written composition. Data obtained from the essays gave more meaningful information concerning the actual performance level of the superior students than the standardized test results. While the superior students committed fewer technical errors in their essays than one would have predicted from test results they met considerable difficulty in writing clearly. All the superior students had not grasped the essential nature of sentence structure, paragraph and essay organization, and punctuation usage, although they were well versed in the rules of usage.

ANALYSES OF CONTENT OF NINTH AND TENTH GRADE ESSAYS

The directions for the writing of the essays in the ninth and tenth grades were given above. Before detailed analyses could begin several research assistants and the writer read random samples of the essays with the following questions in mind. Are there any common patterns in these essays? If so, what are they? What kinds of deviations from common patterns appear, and how frequently? In what ways do these students see themselves? What approaches are commonly used in responding to the questions and what exceptions occur? Were the subjects appreciative or critical of themselves and their environments? Did they describe themselves and others? Which areas were most frequently mentioned or omitted?

At the conclusion of the preliminary reading the fourteen categories listed below were devised to classify the matters discussed in the essays. They served as common bases for comparison of the materials in the ninth and tenth grade essays.

1. Adult Members of Immediate Family. This category includes mention of all the adult members of the student family, namely, parents, grandparents, aunts and uncles, etc. This category will be referred to by the general title of Family in the tables that appear later.

2. Specific Adults. This category includes adults other than parents, relatives, teachers, counselors, and religious advisers. It is a very restricted category and perhaps for this reason the frequency of mention was very low.

3. Siblings. The fact that a few of the subjects were the only child may have influenced the number of "no mentions" which appear in the tables.

4. Peers (friends). In this category analysis was made of mention of the subjects' classmates, playmates, and friends.

5. Local Community Background. In this category mention of the community where the subject lived, or had lived, was considered. It could be a town, a city, or a rural area. It is noted in tables as "Community."

6. Society in General. This category included comments about the world, the United States, and about problems of people, peace, etc. It will be named in the tables as "Society."

7. Religion. This category included comments about religion in general and about specific churches, ministers, church groups and activities, and about the influence of religious principles in the student's life.

8. Schools and Teachers. Comments about schools, subject matter, and teachers were placed in this group. Classmates and school activities were not considered here because they appeared in the fourth and tenth categories.

9. Desire to Learn. This category included the expressed desire to complete one's education, to go to college, and any kind of reference to the pleasure or dislike of studying and learning.

10. Activities. All kinds of activities in which the subject was engaged were considered here. Sports and hobbies as well as intellectual, musical, and manual activities were noted.

11. Choice of Career. This category included the mention of preference for specific careers, the reasons for choosing the careers, and the various alternatives expressed by the subject, as well as concern about his lack of decision.

12. Perception of Self. Included in this category were all kinds of personal reference, and comments about the subject's own feelings about personal qualities and defects.

13. Enthusiasm for Living. This category included any expression of feelings about life in general. It was a difficult category because it was necessary to infer from statements what his thoughts about life were.

14. Reaction to Writing of Essay. In this category any kind of reference to the actual task of writing a specific essay at the Laboratory was noted.

Procedures for Analyses

Each of the counselors read each essay to get a general impression of the content. He then read it a second time and placed on a sheet, which contained the student's name and a list of the categories, one figure to indicate his judgment about the content of that area. In doing so he used the following symbols for each category.

0 - no mention
1 - mentioned in descriptive terms only
2 - critical negative statements
3 - positive or appreciative statements
2 and 3 - both positive and negative statements

(The only exception to the scale was in category No. 11, Choice of Career. In this area the 2 was used to indicate expressions of uncertainty.)

It should be noted that there is no implication of excellence in the use of the numbers. They are simply shorthand symbols to which no scale could be applied. In general they indicate only presence or absence of mention as observed by the reader.

Three graduate students did the categorization and each essay was read independently by at least two of them. In general practice they were paired for the reading of the essays. Counselor number 1 was paired with counselor number 2 thirty-eight times; counselor number 1 was paired with counselor 3 thirty-four times; and counselor number 2 was paired with counselor 3 thirty-four times. Counselor 1 read a total of 68 essays; counselor 2 a total of 66; and counselor 3 read 75.

The reading process involved two phases. The first might be described as a training period during which the counselors acted as a committee. They discussed some essays and the categorization of the items in efforts to find a common frame of reference. During the reading period each counselor worked alone but read the same essays as one of the others. Once a week they met and checked their classifications for the essays read during the period and decided on the final classifications.

In 5,712 categorizations made by the procedures described above, 875 disagreements were found. (The percentage of agreement was 84.7, which appears to be a respectable level for materials of this kind.) Of the 875 differences, 682 occurred in descriptions of critical and appreciative behaviors. The greatest number of differences among readers appeared in the desire to learn, perception of self, and enthusiasm for living categories. Such differences may be due to difficulties in definition of terms and the breadth of the categories.

Results

If each of the 100 ninth and 100 tenth grade students who made statements fell into each of the fourteen categories there could have been 2,800 entries. In Table 3 it will be seen that almost half of the students in each grade did not make statements that could be placed in any category. Of those that did appear almost one-third were positive. Approximately one in ten references made were merely descriptive, and nearly six percent contained mixed positive and negative statements. Only 2.1 percent of the ninth grade and 5.6 percent of the tenth grade statements were negative or critical. In general the figures suggest that when students do choose to make comments they tend to be either positive or merely descriptive. Occurrence of negative or critical statements are so rare that they may be of considerable significance to counselors.

Table 3. Summary of 2800 Categorizations Obtained from One Hundred Ninth and One Hundred Tenth Grade Essays

Classification of Comments	Number and Percents in Ninth Grade	%	Tenth Grade	%	Total	Percent
Descriptive only	141	10.1	194	13.9	335	11.9
Negative	29	2.1	79	5.6	108	3.9
Positive	498	35.5	418	29.8	916	32.7
Mixed positive and negative	73	5.2	89	6.4	162	5.8
Not mentioned	659	47.1	620	44.3	1279	45.7
	1400	100.0	1400	100.0	2800	100.0

In the table the general summary for all categories is presented. A breakdown by specific categories (note exception of career choice category) is presented in Table 4. Since the number for each grade was 100 the figures in the Table may be considered as percentages. It will be observed, for example, that the readers found reference to adult members of the family omitted in only one and eleven percent of the ninth and tenth grade essays. This might be contrasted with the fact that 80 and 85 percent of ninth and tenth graders omitted any comment on society in general. Reversing the figures in the first and second columns of the table one can see that 55 percent of ninth graders mentioned religion but only 34 percent of the tenth graders did so.

Table 4. Percent of Ninth and Tenth Grade Superior Students Whose Essays were Judged to Contain References to Categories to the Extent Indicated

	No Mention		Descriptive only		Negative		Positive		Mixed	
	9	10	9	10	9	10	9	10	9	10
Adult members of family	1	11	3	16	1	5	84	59	11	9
Specific adults	76	90	3	4	3	2	18	4	0	0
Siblings	47	54	11	19	4	5	32	18	6	4
Peers	26	32	26	24	2	4	37	30	9	10
Community background	61	56	18	31	2	1	16	7	3	5
Society in general	80	85	5	8	1	2	13	4	1	1
Religion	45	64	10	2	0	3	44	27	1	4
Schools and teachers	16	18	5	24	0	5	64	52	15	11
Desire to learn	39	29	21	18	0	6	40	45	0	2
Activities	42	30	14	10	0	0	43	59	1	1
*Choice of career	55	25	14	22	11	29	20	24	0	0
Perception of self	20	5	0	2	4	4	54	48	22	41
Enthusiasm for living	57	37	9	7	0	4	31	51	3	1
Reaction to writing of essay	96	84	2	7	1	9	2	0	1	0

*The items for this category for those who mentioned it should be read as follows. Eleven percent of ninth and 29 percent of tenth graders expressed uncertainty about the occupation named. Twenty percent of ninth and 24 percent of tenth graders expressed enthusiasm for the occupation named.

The relative attention paid to the various categories and the general tenor of the comments by ninth and tenth graders in their essays may be observed in Table 5. Despite the differences in directions given to the students at the different grade levels it will be seen that the percentages of ninth and tenth graders whose essays were judged to contain comments about the categories remain very much alike even though there are small differences in the rank order. The categories in which ninth grade response exceeds those of tenth graders by

Table 5. Rank Order Percentages of Ninth and Tenth Grade Subjects Whose Essays were Judged to Contain References to Categories Indicated

	Grade 9			Grade 10	
Percentages of Mention		Most Frequent Kind of Response	Percentages of Mention		Most Frequent Kind of Response
99	Adult members of family	Positive	95	Perception of self	Positive
84	Schools and teachers	Positive	89	Adult members of family	Positive
80	Perception of self	Positive	82	Schools and teachers	Positive
74	Peers	Positive	71	Desire to learn	Positive
61	Desire to learn	Positive	70	Activities	Positive
58	Activities	Positive	68	Peers	Positive
55	Religion	Positive	73	Enthusiasm for living	Positive
53	Siblings	Positive	46	Siblings	Descriptive
43	Enthusiasm for living	Positive	44	Community	Descriptive
39	Community	Descriptive	36	Religion	Positive
24	Specific adults	Positive	16	Reaction to writing essay	Negative
20	Society in general	Positive	15	Society in general	Descriptive
6	Reaction to writing essay	Positive or Descriptive	10	Specific adults	Positive or Descriptive

more than 10 percentage points were adult members of the family, religion, and specific adults. Tenth grade exceed ninth grade mentions by more than 10 percent in the categories of self-perception, activities, and enthusiasm for living. (Seventy-five of the tenth graders mentioned choice of career while only 45 percent of ninth graders did so.) It will be observed that in 19 of the 26 entries in Table 5 the majority of the comments were of a positive nature, in 4 the greatest number were descriptive only, in 2 categories equal numbers of descriptive and positive statements appeared, and in only one area, reaction to writing the tenth grade essay, the comments were negative.

Sample Ninth Grade Essay
(not edited)

You will be given this hour to write an essay about yourself entitled, "The Dominant Forces That Have Directed My Life." Perhaps it would help if you keep the following questions in mind while you write.

1. In what ways have certain persons influenced my life most?
2. What other factors have caused me to be the person I am?
3. How will the above two factors aid or hinder me in attaining what I hope to become?

"The Dominant Forces That Have Directed My Life."

"I believe that my personality has been affected perhaps most strongly by only a few occurrences. Of course, formative changes of less scope which were not so obvious determined the main traits of character which I possess, but since I can recall none of them distinctly, their existence and effect will have to be taken for granted.

I have always been more quiet and restrained than most of my friends. But this characteristic was developed fully at the beginning of my sixth grade in school. It was at this time that I was required (by a change in zoning) to transfer to a different school than the one I had previously attended. That I should thus become even more restrained can easily be seen in view of my past personality. At first, after this change, I seemed to have tried to gain normal friendship with my new classmates. However, this attempt ended, after a good fraction of the year, in a state of no further forward progress. Whether this was due to my own failure to push on, I don't know. Probably, the cause was a combination of recessive tendencies ascerting themselves, and new friendship.

My main friend at the new school was one who, as I was, transfered there from the same previous institution that I attended. Although he was somewhat more successful in his unconscious drive towards acceptance than I had been, we became friends. From this time on, both of us seemed to form certain parallel opinions. Undoubtably our companionship was one reason for this, but it was not a case of one leading while the other followed. Rather, our joint contributions served as a force accelerating both of us faster than either would have progressed alone.

The opinions we formed were essentially detrimental to any further great advance towards becoming more social. The frivolity of classmates made us actually have no wish to become completely accepted, if such entailed the prerequisite that we adopt the same type of activities as those whom we might have wished to be good friends of. Technically here, it is almost meaningless to speak of "friend" in the singular. The term must actually apply more to the group of classmates, rather than to individuals. As the situation (the actions and modes of behavior of others) has substantially maintained the same status, and as my own opinions have not been deserted by me, I am not now a staunch member of the group, nor do I desire to be.

My future will certainly be affected by this policy. How depends upon relatively unpredictable factors. That is, I cannot predict them. Whether others will change gradually to a state of more individuality, I do not know. Perhaps the status quo will be maintained. I have no dislike for people, merely disagreement. If the situation does not change, I believe, however, that I will nevertheless get along, as I am not misanthropic. Perhaps I shall choose some profession in which most of my contacts are made with people having similar ideas. At any rate, my future shall be determined greatly by the opinions I have formed from sixth grade to the present."

Sample Tenth Grade Essay (not edited)

The purpose of this assignment is to make you aware of the variety of things that influence us in our "growing up" and in developing plans for our future. Besides being an assignment in writing about a very interesting person, YOU, the autobiography will help you decide "What kind of person am I?" "How did I get that way?" and "What do I hope to become?" Keep these three questions in mind and write freely about yourself. Include anything that you feel helped make you what you are. You will have about an hour to write.

> "I am not perfect. I am quite intelligent, but I am also quick-tempered, shy, and impatient. I will now pick my personality apart and examine it.
>
> My intelligence contributes much to my personality. I have been repeatedly told that my intelligence is a gift of God, so I don't brag too much about it, at least, I don't think I do. The fact that I get my work done quickly, and so have been idle many times, should have gotten me into trouble. However, I see no sense in many kinds of troublemaking and am uninterested in any of the rest except for talking, which I do now and then.
>
> One thing that fills my time when I have no schoolwork is daydreaming. My dreams are mostly fiction based on novels, actual events, and even textbooks. Formerly I took revenge on people I didn't like temporarily by having my main character come and knock them around a bit. At first my hero was a mysterious bearded American millionaire. Then I switched to an imaginary foreign country. Before I came into my present phase, I took all actual persons from them. I now have my foreign islands, a group who is going to overthrow the world, and a group of American guerillas in the III World War.
>
> My quick temper gets me into trouble mainly at home, since I rarely, if ever, lose my temper elsewhere. My main targets are my younger brothers and sisters. My main reasons are their not doing their chores on time and their fooling around with my things.
>
> I am very shy, especially in the presence of strangers or my elders. At social gatherings and other chance meetings I am excessively shy with girls, even with those I know. In class or other places like it, I can talk fairly easily with them, but I cannot initiate a conversation.

I have few close friendships and those that I have are mainly caused by proximity in class and with at least one, somewhat similar interests.

I am impatient and when someone is late I worry excessively. I walk around, mutter to myself, and generally make a hard time of it.

I would like to change this, particularly my shyness with girls, as certain signs in the past year have produced in me a desire and need for marriage. I am trying to conquer my quick temper and though it's a slow, hard battle, I think I'm winning.

I often doubt. I doubt almost anything I have ever been taught. I still disbelieve evolution, perhaps more strongly now than last year. I sometimes doubt my religion, however, those doubts don't remain long. Several times I have doubted existence, gravity, knowledge, etc. You name it, I've doubted it.

I don't know where my thirst for knowledge comes from, but I know I like knowledge, any kind of knowledge. I know a little bit about a lot and a lot about a little bit. I like books of all kinds, though some books I consider dull.

Summing up, I am intelligent but some of my other qualities aren't so good.

ANALYSES OF ELEVENTH GRADE ESSAYS

The essays of 100 high school juniors were selected for analysis at random from the 185 which had been written at the Laboratory. The students came from 49 participating schools during the school year 1960-61. The directions given for the writing of the essay appear at the beginning of the chapter. The title of the essay was, "When Should Teenagers be Treated as Adults?"

The formulation of appropriate categories for analysis of this essay resulted from thinking about many questions, suggested both by the topic and a general reading of 30 essays. All were asked in attempts to get an answer to that primary question, "What kinds of things may we learn about a student in this essay?" Questions such as these required answers that would reveal not only what the student said, but how he chose to say it as well. To what extent were these superior students influenced by the structure of the topic presentation? Did they write about the points mentioned in the outline? Was originality of style and comment characteristic of their writing? How concerned were these students about the matter of being treated as adults? Was this concern reflected within the 2 to 5 pages of the essays they wrote? Were attitudes about adults and peers reflected in their writing? Was concern about the future,characteristic of many adolescents,brought into the discussion?

Thirty of the 85 essays not used in this study were read in order to see if categories could be set up that would make analysis of the essays possible. The following eight categories emerged.

1. Extent of personal involvement. All personal references and mention or absence of mention of self were included here. Involvement referred to the extent to which the writer related the topic to himself.

2. Extent to which outline was followed. The classification indicated how closely the writer followed the suggestions for consideration outlined in the topic question. If the writer chose to discuss one of the 4 points mentioned in the outline, his essay was categorized as partially following the outline. No differentiation was made in the results between those discussing only one suggested idea and those discussing as many as three.

3. Acceptance or rejection of criteria stated in topic outline. This category included writers' ideas of when maturity is achieved according to the extent to which they agree or disagree (accept or reject) with the criteria for maturity stated in the topic outline.

4. Originality of thought and/or method. In this category were included all addition of ideas or expansion of topic and uniqueness of style or approach. An idea or example was considered original if it was not stated or suggested by the topic question. This category is referred to in the table as originality.

5. Degree of concern about the topic. This category described the overall tone of the essay as a reflection of the writer's concern about the topic. It was felt that a definite impression of the writer's concern would be given by the quality of his comments. Concern refers to interest in thinking about the topic whether as a developmental problem or as a source of personal conflict. This category is referred to in the table as degree of concern.

6. Attitude toward adults. All comments about specific adults and the adult world in general were placed in this category. In some cases where attitudinal content was mixed, the essay was rated according to that attitude judged most dominant in terms of overall emphasis and amount of discussion.

7. Attitude toward teenagers. All comments about teenagers as a group and about specific friends and acquaintances were classified in this area.

8. Attitude toward the future. This category included all comments about the future or any discussion projected into the future by means of mention of outcomes or eventual circumstances. This category was used on the basis of its possible inclusion in the essay rather than on actual evidence of these attitudes in the essays read initially. For this reason the frequency was expected to be very low.

For each of the 8 categories 3 or 4 statements descriptive of what students' wrote were devised. These statements follow.

Extent of Personal Involvement

A. No mention of self (discussed topic in terms of teenagers in general, gave examples of their behavior or ideas rather than own, did not personalize discussion).

B. Incidental mention of self (included self in topic in terms of personal opinion and/or example and explanation).

C. Frequent mention of self (discussed topic primarily in terms of self, on a personal or confidential level).

Extent to Which Topic Outline was Followed

A. Outline completely followed (remarks all directed toward questions suggested, specific answers given to each question).

B. Outline partially followed (some questions selected for discussion, perhaps some additional ideas included along with these).

C. Outline not followed (use of a different approach and organization for discussion of the topic, concepts stated, original terms defined, etc.).

Acceptance or Rejection of Criteria Stated in Topic Outline

A. Complete acceptance (All suggested criteria supported as reasonable determinants of maturity, agreement with societal standards commonly used as measures of maturity).

B. Selection of items (certain items discussed in terms of both acceptance and rejection with qualifying statements given for explanation of point-of-view and limitations of criteria as stated).

C. Complete rejection (all suggested criteria rejected as means of determining maturity, concept of individual differences or unsuitability of external standards given as support of this rejection).

Originality of Thought and/or Method

A. Absent (used ideas suggested by the topic question, added no original examples or illustrations, evidenced no thoughts out of the range of the obvious for this discussion, approach generally non-imaginative and non-individual).

B. Present (used examples, illustrations which were not immediately suggested by the topic question, defined terms, drew analogies, discussed cause and effect relationships, used descriptive or picturesque language, exhibited original style of organization and/or language, approach indicated individual thinking and imagination).

C. Present to high degree (used technique of story-telling or fantasy in approach to topic, extreme degree of originality and imaginative thinking in presentation of whole discussion of topic).

Degree of Concern About the Topic

A. Indefinite (no definite impression of concern and commitment given to reader).

B. Casual (topic taken "lightly," superficial discussion of points, little apparent interest, extremely short length, humorous or sarcastic approach, etc.)

C. Serious (evident concern in discussion of topic, commitment to definite point-of-view discussed seriously, detailed presentation of arguments, persuasive, supportive statements, or extremely long in length).

D. Critical (comments in general gave criticism of topic as unimportant, irrelevant, or unsuitable for discussion. Unwillingness to discuss topic because of objections to it!).

Attitude Toward Adults

A. Objective (did not express personal feelings about adults, discussed them neutrally in terms of their relationships with teenagers, their standards and expectations).

Attitude Toward Teenagers

A. Objective (did not express personal feelings about teenagers, discussed them neutrally in terms of their behavior, problems, etc.).

B. Criticism (was critical of teenagers' behavior, blamed them for immaturity, chastised them for not assuming an adult role, or pointed out why they are not treated like adults).

C. Defense (defended teenagers, stated their dilemma, pleaded case of great maturity of teens, identified with dual child-adult role of a teenager).

D. Sympathy (felt sympathy for situation of teenagers, stated an understanding of their problems, "One minute you act like a baby, the next minute you are grown up").

Attitude Toward Future

A. No mention or no impressions of attitude evident (did not refer to future).

B. Generally optimistic (made positive statements about the future and its outcomes, hopeful, "things will be better" point of view).

C. Generally pessimistic (made negative statements about the future and its outcomes, probable circumstances discussed negatively).

A copy of the sheet which was used to record the classifications of the reader of the essay appears below. It should be noted that the letters which appear do not indicate any degree of excellence. They are simply headings to indicate descriptions of what the reader observed.

An important step in this process of planning essay analysis was testing the reliability of the instrument. If there could be substantial agreement between 3 separate judges analyzing the same documents, it could be assumed that the categories and descriptions that had been devised were appropriate and adequate, and that confidence could be placed in the reliability of the instrument. The following procedures were employed to determine observer reliability. Twenty-five essays were read by 3 judges. All were working toward the master's degree in Guidance and Counseling. Training sessions were held at which essays of subjects drawn from the same population as those in this study were read, descriptions compared, and categories discussed and clarified. This discussion led to slight revision of the defining characteristics of categories No. 2 and No. 4 and better understanding of the total group of categories as applied to the essays. After meeting together, the 3 judges independently read and categorized the materials in 25 essays drawn from the total sample of 100 to be used in this study.

ESSAY ANALYSIS RECORD

Student's Name ________________

School ________________

M or F ________________

Category	A	B	C	D
1. Extent of personal involvement				
2. Extent to which outline is followed				
3. Acceptance or rejection of criteria stated in topic outline				
4. Addition of ideas, originality				
5. Degree of concern about the topic				
6. Attitude toward adults				
7. Attitude toward teenagers				
8. Attitude toward future				

Comments: (Add examples and note any unusual aspects of the essay.)

Overall agreement between judges on all essays was 97.5 percent. The high percent of agreement was due to the following factors:

1. the categories were devised through the content analysis method and actually represented the expected content of the essays;
2. the categories were broad enough to include the possibilities for probable variation yet specific enough to make categorization rather precise; and,
3. the training period brought the judges to a common understanding and unified approach in their analysis.

The listing below of the percent of agreement for each category shows that for 4 of the 8 categories (1, 4, 6 and 8) the essays were all read in the same way by the 3 judges. The lowest percentage of agreement in category 3 appeared to be due to differences in interpretation of the writers' views of the criteria of maturity stated in the topic outline. The remaining 3 categories (2, 5 and 7) were read with 96 percent agreement. One unusual essay produced disagreement in each of these categories. It appeared from the study of reader reports that, when they follow the procedures utilized here, they can reach a high level of agreement. The results do not, however, indicate that the readers' analysis are necessarily valid.

Table 6. Percent of Agreement Between Three Judges on Categorization of Responses in Twenty-five Essays

Category	Percent of Agreement
1. Extent of personal involvement	100
2. Extent to which outline was followed	96
3. Acceptance or rejection of criteria	92
4. Originality	100
5. Degree of concern about topic	96
6. Attitude toward adults	100
7. Attitude toward teenagers	96
8. Attitude toward future	100

Findings

The 100 essays were analyzed to discover (a) the typical pattern of response in terms of approach and content in the writing of the essay, (b) the atypical patterns of response in approach and content, and (c) characteristic attitudes of superior students about achieving adulthood, maturity, the adult world, peers, the future, and personal concerns. The frequency of particular responses in the various categories gives form to whatever patterns have emerged. Discussion of the analysis and ratings for each of the 8 categories make these patterns clear and identify the attitudes expressed and reflected in the essays.

The sample of 100 essays was read to determine the extent to which the subjects became involved personally in the topic discussed. This personal involvement was indicated by quality and quantity of reference to self. From Table 7 it will be seen that the most common pattern in these essays was to make only incidental mention of self while discussing the question of when teenagers should be treated as adults. Only 20 percent of the writers did not refer to themselves but 3 times as many boys as girls did not do so. The remaining 3 percent became involved in the topic to the extent that they related the problem directly to themselves and discussed personal concerns in the writing of this essay. It appears, then, that in considering the question, "When should teenagers be treated as adults?", these students tended to write impersonally with mention of themselves primarily in terms of generalized personal opinion about the whole topic. The question, as presented to them, asked for personal opinion about the "answer" and would thus encourage personal reference. The absence of greater personal involvement in most of these essays may in part be due to techniques of expository writing learned in school which tends to emphasize great attention to exposition of the topic with personal references used only when appropriate.

Table 7. Extent of Personal Involvement Evident in Essays on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. No mention of self	10	30	20
B. Incidental mention of self	86	68	77
C. Frequent mention of self	4	2	3

Study of Table 8 reveals that students usually followed the topic outline in part and based their discussions on one to 3 of the ideas suggested by the outline. The same number of boys and girls, almost a third of the total sample, followed the topic outline completely and discussed all of the points presented to them for

Table 8. Extent to which Outline was Followed in Essays on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. Outline completely followed	28	28	28
B. Outline partially followed	62	58	60
C. Outline not followed	10	14	12

consideration. Twelve percent of the writers disregarded the organization and ideas given in the topic question. A majority of students were selective about the ideas they discussed and were discriminating in their response to the total stimulus presented.

Analysis of 100 essays for the third category yielded ratings in only 2 of the 3 descriptive classifications possible. From Table 9 it may be seen that none of the subjects expressed complete acceptance of the criteria for maturity stated in the topic outline. These criteria were presented in question form for the reaction of the writers, and the most frequent pattern of response was selection of items for discussion with little non-critical acceptance of any criterion. The remaining subjects, 24 percent of the total group, completely rejected all of the stated criteria as unacceptable. Boys and girls were in close agreement in their views on suitable means of determining maturity in teenagers. The pattern of typical response indicates that the suggested "black and white" concepts of maturity were not held by these students. Their thinking about maturity and adulthood was critical. The awareness of exceptions to the criteria and the complete insistence on these exceptions are characteristic of this critical thinking.

Table 9. Acceptance or Rejection of Criteria in Essays on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. Complete acceptance	0	0	0
B. Selection of items	78	74	76
C. Complete rejection	22	26	24

Originality was conceived as a broad area, somewhat related to the second category in the extent to which the topic outline was followed, yet aimed more at discovering the unique characteristics of each writer's performance. The category provided for classification

Table 10. Originality Evident in Essays on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. Absent	42	44	43
B. Present	52	54	53
C. Present to high degree	6	2	4

of original style or method as well as for addition of ideas not actually suggested by the topic question. Originality, then, represents those ideas and ways of presenting them discussed in the essay as a result of the individual's personal frame of reference and not specifically indicated in the topic question. The figures in Table 10 suggest that slightly more than half of the subjects exhibited some originality in their essays and did not confine their remarks to the ideas immediately apparent upon reading the topic question. Again, the similarity of performance between boys and girls should be noted. Only 4 percent of the total subjects showed a high degree of originality (see example later). Forty-three percent of the students added nothing to the essay that had not been stated or suggested in the topic question.

The results of the analysis of essays for the fifth category (the degree of concern about the topic) indicated that most of the writers produced thoughtful, serious compositions about the question of adulthood and maturity. Because of the impressionistic nature of analysis of this category, the indefinite category proved useful in classifying the 15 percent of essays in which there was no clearcut evidence of the intent of the writer. Rather than describing them inaccurately, the investigator grouped those essays into one class. The girls appeared to be more committed to the topic as a group. Eighty-six percent of them took the topic seriously and evidenced some conviction and interest in their writing while only 8 percent gave no definite impression of committment. Only 2 percent wrote casually or lightly about the topic. The boys, however, were represented by 22 percent in the indefinite category as compared to 8 percent of the girls. Only 3 percent of the total group were critical of the topic given. It will be shown later that the comments of these students on maturity, and their concern about achieving adulthood, reflected a great deal of thought and serious consideration.

Table 11. Degree of Concern Evident in Essays on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. Indefinite	8	22	15
B. Casual	2	4	3
C. Serious	86	72	79
D. Critical	4	2	3

Two of the 3 areas of attitude analysis were very similar in form and yielded similar results. It may be seen in Table 12 that the most frequent type of reference to adults was in neutral terms without expression of personal feelings. The more positive attitudes of the last two categories in Table 12 were held by 18 percent of the total sample. Opportunity for airing grievances against the adult world present in this topic was used by only 4 percent of the subjects.

Table 12. Attitudes Toward Adults Evident in Essays on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. Neutral	82	74	78
B. Criticism	2	6	4
C. Reliance	14	16	15
D. Appreciation	2	4	3

The common pattern of objective reference found in the sixth category is evident in Table 13 below. It is interesting to contrast the frequent neutral references to teenagers with the infrequent critical views held by 5 percent of the subjects. The possibilities for more positive views held about teens in the last 2 categories in the table were responded to by only 7 percent of the total group. Thus, of 100 essays analyzed on the topic "When should teenagers be treated as adults?", slightly over one-tenth of the students expressed personal feelings about their peers.

Table 13. Attitudes Toward Teenagers Evident in Essay on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. Neutral	84	90	87
B. Criticism	6	6	6
C. Defense	6	4	5
D. Sympathy	4	0	2

A reminder should be made here that the last category was considered as a possible but not probable response in terms of the ideas suggested in the topic outline. References to the future may be seen in Table 14 to form a frequency pattern which includes nearly all of the subjects. It may be seen that none of the 100 subjects expressed any negative feelings about circumstances or outlook toward

Table 14. Attitudes Toward the Future Evident in Essays on "When Teenagers Should be Treated as Adults"

Classifications	Percent Girls (N = 50)	Percent Boys (N = 50)	Percent Total (N = 100)
A. No mention	98	98	98
B. Generally optimistic	2	2	2
C. Generally pessimistic	0	0	0

the future, and an equal number of boys and girls were represented in the small percentage concerned with the future in a positive way.

Illustrative Writing

A look at views expressed by specific subjects of this study will serve to illustrate the results of analyses of 100 essays entitled "When should teenagers be treated as adults?" Excerpts from representative essays are presented here to illustrate attitudes of the superior students toward maturity, achieving adulthood, adults, peers, the future, and personal concerns. Selection of excerpts was designed to illustrate samples of typical and atypical content. More attention has been given to the fourth, sixth and seventh categories because these areas (originality, attitudes toward adults, and attitudes toward teenagers) were given greatest emphasis by the subjects in terms of quantity of remarks and focus of discussion. All original spellings and constructions of the students have been retained.

Extent of Personal Involvement

The most frequent pattern of response in this category consisted of comments such as "In my opinion," and "I think," along with other similar phrases identifying the writer as an observer or narrator. Illustrations of actual personal concerns discussed by students are given as representation of the 3 percent who personalized discussion of the general topic to a great extent.

> "Opening almost any leading magazine today, one will find a number of articles dealing with 'the poor, confused situation of an American teenager.' Personally, my situation is not confusing. Perhaps in the present day I would be considered maladjusted, since according to the 'experts' we teens are supposed to

> be at that age when we are experiencing the difficulties of acclimating ourselves to society.
>
> I do not consider myself an adult - that is not completely and not yet. In all frankness, I do consider myself more of an adult than many of my compatriots, however I have not had all the experiences which go into the composition of sound judgment and reasonable decision-making capacity. I am treated like an adult when I act like an adult and when...I act like a child or a 'teen' I am treated accordingly."
>
> (Girl #8095)

> "This brings up one of my pet peeves. Being only fifteen, I'm treated by my mother as being too young to accept responsibility. I feel that I'm old enough to realize more responsibility. She believes that age is the rule that sets responsibility...The people at the store respect me and trust me because they think that I can accept responsibility. They showed me this when they left the store in my charge when they went up North."
>
> (Boy #8052)

The content of these 2 essays is very different yet they were both written on the same topic "When should teenagers be treated as adults?" and they both reflect very personal concerns of these 2 individuals.

Acceptance or Rejection of Criteria

Student's views on appropriate determinants of maturity were found to be very similar in the essays. The most common pattern of view concerning the criteria for determining maturity was that of qualified acceptance. While none of the subjects accepted these criteria completely, 24 percent considered them all unacceptable. Samples of arguments frequently given for partial acceptance and complete rejection of the given criteria are as follows:

> "I think that teenagers should be treated as adults when they are old enough to accept responsibilities... An adult is not more, in a sense, than accepting responsibilities. When you're an adult, you are responsible for the upkeep of your house, feeding your children, support of you and your wife, and being a working cog in society.
>
> (Boy #8052)

> "By the time we are out of high school we should be mature enough to lead an adult life."
>
> (Boy #8094)

The students expressing rejection of the criteria for maturity suggested by the topic outline were generally much more emphatic about their views than those accepting one or more criterion, such as those quoted above. The next statements are in agreement that a definite time for maturity cannot be predetermined.

> "Maturity doesn't come with a high school diploma, when you have a regular job and are self-supporting or when you reach a certain age. Maturity comes with experience, experience with people and experience with things."
>
> (Girl #8150)

> "There is no age for maturity. There is no date at which one can say, 'Now I am a man!' Others may, for the purposes of generalities, classify maturity, but even there is the recognition that the generalization does not always apply."
>
> (Girl #8095)

> "I think people mature at different rates and... personal experiences determine the maturity of the individual."
>
> (Boy #8093)

> "There is no definite yardstick for maturity. Every person is a separate and complete entity and develops at a different rate than any other person."
>
> (Boy #8053)

Originality

Addition of ideas and expansion of topic, characteristic of originality (as used in this study), were well exhibited by the subjects in their discussions of the concept of maturity. The comments about maturity given here are especially appropriate since so many students chose to focus upon it and discuss it in great detail. These comments show a variety of views about the meaning of maturity. These views might be considered clues to personal values of the writers and are thus worthy of notice. Originality, presented here in terms of individual thought and belief, was an important characteristic of slightly over half of the 100 essays analyzed. As evidenced in the excerpts following, the product of this originality provides answers to the questions, "What is maturity?" and "What is a mature individual?"

> "Think of all the things which come to your mind when you hear the word maturity: marriage, clear thinking, good decisions, common horse sense, taking your place in the world."
>
> (Boy #8017)

> "Qualities of maturity: awareness, sense of humor, a sincere interest in other people, pride in humility, love for all others."
>
> (Girls #8150)

> "Maturity isn't something that can be assigned to everyone like a license number or a chair in a theater."
>
> (Boy #8016)

> "Maturity is a regulation of will by intellect. Its tissue is determined by convention."
>
> (Boy #8162)

> "The mature individual: he makes his own decisions, but accepts the suggestions of others when they are given. He is open-minded and fair. He has definite opinions about life in general and is not afraid to express them. He respects the rights of others and does not feel that the world owes him a living."
>
> (Girl #8025)

> "Consistency is the major objective of those seeking maturity. The first step toward maturity is, then, the construction of goals and convictions. The second step is acquiring ability and courage to meet these goals and satisfy these beliefs."
>
> (Boy #8137)

> "When a young person has to remain at home and help a sick relative every hour that she is not in school, she may mature more rapidly than her counterpart who is with other people her own age who just run around and do very little."
>
> (Boy #8039)

> "Any boy who participates in a school sport and earns a letter not only by playing good but also showing respect for sportsmanship and other partici-

> pants, and then applies this to everyday life is very worthy of the respect of an adult."
>
> (Boy #8058)

Attitudes Toward Adults and Attitudes Toward Teenagers

The greatest yield of attitudinal content in the essays analyzed was in the 2 categories named above. They are presented together because in many instances the subjects expressed their views about both groups at the same time. The relationship between adults and teenagers is an almost inevitable discussion point because of the specific topic question "When should teenagers be treated as adults?" The common patterns of objective reference to adults and teens discussed previously are represented in the excerpts given, but more attention has been given to attitudes less frequently stated but more individualistic. These are attitudes which must be identified and understood for they are indicators of students' feelings toward the people about them.

> "A teen should be given more and more freedom and independence until eventually he should be free from all strings which tie him down and make him feel an inferior human being whose actions must be governed more than others."
>
> (Boy #8063)

> "Control over a teenage child should not be relinquished at too early an age...the child will feel lost; without control there are too many decisions to make. Adults have years of experience plus all the rules they learned from their childhood to guide them in making decisions, but a teenager, if set loose at too early an age, lacks guides to what society expects of him and to what is best for him."
>
> (Girl #8049)

> "As far as I'm concerned, I want to be treated as an adult when my parents and others feel I'm ready, not when I reach a certain age and state laws say I should now be considered an adult."
>
> (Girl #8112)

> "Proof of maturity can be seen easily by the parents, who can find out just how much they can trust their children by giving them more chances to prove that they are old enough to be adults. The parents still have to expect them to act like other kids their age

> once in a while because the young people of today want to be popular and fit in with the crowd."
>
> (Boy #8043)

> "Neither experts nor society nor law determine maturity. That is for the parents to fulfill...."
>
> (Girl #8133)

In Defense of Teenagers

> "Because a person is a certain age, teenage for example, should not make him suspect of all the psychological maladies and weaknesses known to society."
>
> (Boy #8020)

> "It isn't always the teenagers' fault. Maybe the parents are too strict on their youngsters."
>
> (Girl #8134)

> "The typical teenagers is bound to rebel or hold a grudge some way if he or she is pushed around and used as a type of puppet that can't think, talk or do anything for himself."
>
> (Girl #8117)

> "I think children are pushed into adulthood at a fast pace and they aren't having enough enjoyment of their childhood. Young children today desire things of adulthood at times when, in relationship to times past, they would still be content to play with toys, games, etc. This is probably a result of the generally accelerated state of the world today. To keep up with the fast pace children must grow up rapidly. I think the pace should be slower with more time for relaxation and time to enjoy life."
>
> (Girl #8074)

> "They include as adults many people who are actually children and they exclude many mature responsible young people merely because of age."
>
> (Boy #8045)

> "Many adults spend their money foolishly. Many grown people who are considered adults act like children who can't have their own way."
>
> (Boy #8051)

> "Very few teenagers are mature. Most of them think only of themselves and lose their tempers when some of their trivial wants are refused them. Some can't be relied on. They aren't at all dependable."
> (Boy #8038)

> "I have seen many seniors act so childish. Many of them try to show off or think they are something special by smoking and drinking."
> (Boy #8043)

Given here is one example of the 2 percent of positive response about the future found in this analysis of 100 essays.

> "My general ideas toward maturity are that a person gets his start in high school. After 4 years of college or whatever a person does, he should be nearly or all the way developed. A few odds and ends may be left but the years after that should remove defects."
> (Boy #8053)

Summary

One hundred essays, written during students' third year visit to the Research and Guidance Laboratory for Superior Students were analyzed to identify:

a. typical approaches and content in the writing of the essay;
b. atypical or individualistic approaches and content; and,
c. students' attitudes toward achieving adulthood, maturity, adults, peers, the future, and personal concerns.

The following pattern of approach and content was found to be most characteristic or typical. In considering the question "When should teenagers be treated as adults?", these students tended to write impersonally with only occasional mention of their personal opinions. They followed the outline of suggested criteria for maturity partially by selecting certain of the items for argument or agreement. Most of the subjects did base their discussions on at least one of the suggested criteria but they commonly expressed qualified agreement with them. No one accepted all of the criteria completely although some rejected all of them.

More than half of the students added ideas to those given from the topic outline. Many of their additions were items of personal opinion about the concept of maturity. Almost as many students restricted their comments to the points suggested by presentation of the topic question. Because the subjects who did not exhibit originality were nearly as many as subjects who did, no pattern of typical content can be accurately stated but the greater tendency would seem to be toward some originality in discussion. Most of the boys and girls seemed to be genuinely concerned about the question of achieving adulthood. They were serious in their discussion of the points on which they chose to focus. The length and depth of arguments presented indicated previous thought and perhaps instruction on the matter of maturity.

A neutral view of adults and society in general was characteristic. Standards set by institutions and by parents seemed to be taken rather matter-of-factly with little evidence of conflicts with the adult world apparent in the essays. Neutral views of their peer groups appeared in the discussions of when privileges should be granted, and responsibilities assumed. Attitudes toward adults and teenagers alike tended to be accepting and rational.

Implications

Underlying the purposes and procedures of this study is the basic principle of guidance concerning the necessity of a comprehensive collection of longitudinal data about each counselee. The words comprehensive and longitudinal define the desirable scope of these data as continuous or developmental and varied or broad. These 5 minimum criteria for effective and efficient data collection have been recommended by the writer and are presented here as guides for further practice:

1. Any datum about an individual that assists in the understanding of his behavior must be given due consideration.
2. Any datum about an individual that is to be used in his guidance must be appraised accurately, fully and economically.
3. The culture in which the individual is reared must be thoroughly examined.
4. Longitudinal data must be used in the study of the individual.
5. Conceptualization must be continuous as each separately evaluated datum is added in the study of the individual.

Implications of this study for guidance may be realized through consideration of items one and 2 above. The use of many forms of personal documents available to school staff members could assist counselors to better understand personal behavior and thought characteristics of youngsters _from their own frame of reference or point of view_. Effective and efficient use cannot be made of such documents unless they are thoroughly appraised as counseling tools.

The results of this study have led to certain conclusions about the use of a structured, assigned essay (about a topic of presumed interest and concern to adolescents) as a counseling tool. It may be utilized in terms of: (1) information provided about significant beliefs and attitudes of the individual student, (2) information provided about the individual in comparison to his group, and (3) addition to and substantiation or contradiction of the various other types of data collected.

Eleventh Grade Essay
(not edited)

"I really don't know exactly where to begin. First of all, I definitely do not believe that there is one age at which you are magically transformed into an adult. You cannot say "I am ____ yrs old, therefore I am mature." It does not logically follow. Obviously maturity is a result of not only ones own character, but also the myriad influences that have figured into ones life.

I have always believed that a person has a sometimes conscious, sometimes unconscious tendency to try to live up to what is expected of him.

Perhaps I speak only for myself - I don't know. However, since I was a child, my parents have been relatively lax as far as discipline goes. They have expected certain things of me, yet they never laid down definite rules by which I was forced to abide. That is, I never had to everything snap-snap-snap. I was always trusted to make my own decisions concerning my affairs, tho, on occasion when I lost sight of my elusive common sense, I could feel the pressure of their guiding hands. I was treated much as an adult - tho taught respect for my elders, and as a consequence, developed a sort of conversational-equality relationship. Now I pretty much set my own rules - regarding curfews, studying etc. & all other "typically" teenage points of conflict that might arise, yet all these rules are generally governed by an inherent sense of duty toward my parents & a large dose of self-confidance, the result of "years" of experience. Thus I can say, I have gained & lost in many respects.

I have gained independance, yet lost some security. I have perhaps found more strength in myself as an individual, yet I have often felt the nagging pains of regret & indicision, with only myself to blame. What I gained in a friend (of my parents) I lost as confidantes & advisors.

This may seem a bit off the subject, but I just wanted to go into it, because having had a remarkably happy life, (all of my 16 yrs), I can feel this is an excellant way to raise a child - that the sooner he is treated as an adult, the sooner he begins to behave like one. I realize, however, that this is largely an individual proposition & varies greatly from person to person. Maturity must be practiced, coming only with yrs of "mellowing" & one should consider the "adulthood" of a "teenager" in that light.

ANALYSES OF THE SENIOR ESSAYS

As indicated previously, the senior essay used at the Laboratory required the students to write on the topic, "The Future as I See It and My Place in That Future." The students were given about 45 minutes to write on the subject but if any of them wanted more time they were not held to the time limit.

At the time this study was begun, 126 (64 girls and 62 boys) of the total senior population of 138 students had visited the Laboratory and had written the essay. Preliminary and independent reading of the essays by 2 graduate students in guidance indicated that the written responses tended to fall into categories. Consideration of such categories by the 2 readers and the writer who had also read a sample of the essays resulted in the setting up of the following categories which were used in analyses of the essays.

1. Scope of view
2. Forces that shape the future
3. Condition of the world in the future
4. Definiteness of personal plans
5. Interest in formal education beyond high school
6. Concept of vocational roles and range of vocational aspirations
7. Desire for family membership
8. Unique responses and particular concerns or problems

1. Scope of view.

In reading of the essays the readers were concerned about the extent to which the subjects would write largely about personal goals or whether they would conceive of the future in more global terms. If the latter conception appeared would the subjects see themselves as participants in, and perhaps contributors to, the global development?

The results of the analyses appear in Table 15 below. Four of the students wrote essays which were so extremely personal that they were put in the last classification in the table. More than half of the girls and two-fifths of the boys wove their own plans into the development of society while the remainder were almost equally divided among the polar positions. A study of those who expected to be contributors to development of society is reported later.

Table 15. Scope of View of Subjects Expressed in the Senior Essay

Classification	Girls N	Boys N	Total N
Broad social view with no personal reference	10	18	28
Narrow personal view with neglect of broad social view	10	16	26
Social and personal view related	42	26	68
Extremely narrow personal view	2	2	4
TOTAL	64	62	126

The following quotations from 2 essays typify the point of view, both personal and social in outlook, which was most commonly observed.

> "I do not seek fame. I only look for a chance to make this world just a little bit better to live in. It is my hope that the things I do in the present are preparing me to do a little something for someone in the future."
>
> "Perhaps the greatest challenge and the greatest problems lie in the laps of political and military leaders of the United States and the world....now and in the future, the most important problem will be the establishment of peace. In this future I see myself attempting to do an infinitesimally small part, not only in my profession but in actions as a man and a citizen."

2. Forces that shape the future.

Only 26 of the 126 students addressed themselves specifically to this topic at a level other than struggles between nations. The data do not lend themselves to tabular presentation but the following lists with brief samples of responses indicate some of the thoughts of the subjects who chose to write about such matters.

		Girls	Boys
God:	"Only the good Lord knows for certain"	3	1
Fate:	"Whatever will be, will be" "No one can tell what tomorrow holds"	3	2
Man:	"Man can control his own fate to some extent" "The type of future one has depends on what they do to develop it. The future is *made* by the individual, not dreamed about"	5	4
Government:	"My future will be tied up with the future of the country itself"	1	1
Education:	"College is the key to the future"	5	1

It appears that these students either did not see, in the topic, a challenge to write about forces influencing the future or decided not to discuss a topic that has been debated by philosophers for many centuries.

3. *Condition of the world in the future.*

The diversity of opinion on this matter ranged from extreme optimism (unity of man in a world of peace) to pessimism expressed as the end of the world brought about by bombs to "creeping socialism"

Table 16. Brief Interpretations of Students' Description of the Condition of the World in the Future

Classification	Girls N	Boys N	Total N
"Don't know, uncertain, unclear, one big question mark" world at crossroads	6	7	13
Complex	5	7	12
Challenging, good opportunities	1	5	6
Unity of man in world peace	1	-	1
Confident optimism	6	4	10
War is not inevitable	4	3	7
Troubled times ahead	5	6	11
National struggles and conflict	12	9	21
Creeping socialism	1	-	1
Possible end of the world	-	1	1
Mere continuation of past	-	1	1
No comment on world condition	23	19	42
TOTALS	64	62	126

or a hectic fast life with many problems. A summary of the readers' interpretation of the students' conceptions of the condition of the world in the future is presented in Table 16. It indicates the diversity of opinion that these potential future leaders displayed. Generally the mood is pessimistic rather than optimistic and the number who see the world of the future as a challenging place which will provide good opportunities is very small. Some illustrative comments follow.

> "With all sections of the globe having such separate and definite philosophies of life...and not being able to communicate with each other, they are rushing toward something that no one really realizes and accepts."

> "I do not believe that peace is possible until Communism no longer exists."

> "When I think about the future, all I can see is tension between people and nations."

> "This country is engaged in a cold war now, and I believe that it's going to get tougher."

> "Bright, clear, and sunny. That's how my future looks to me."

Approximately one-third of the subjects indicated that they expected to see progress in technical and human (social) developments within their lifetime and about 30 percent saw themselves as contributing to such progress. Two quotations which illustrate these points of view follow.

> "The future as I see it will be one of great technological advancement. As it seems to me now, automation will play an ever-increasing part in the lives of everyone. Automation, or advancement in any technical field, always brings with it more advancement in mathematics and allied fields."

> "When man in general becomes mature enough to take the place in the world intended for him, major troubles will cease. I feel that it is my obligation to do my part to create a condition like this. I must attempt to be a leader toward a purer democracy, voice my opinions on all subjects, and

> in every way possible try to create a better world. I can do this in many ways. Number one on the list would be to get a good solid education. I should be active in community affairs. I should work to see that competent men are put in public office."

It became apparent to the readers that a number of students felt in some way that everyone had a "duty" toward progress in one way or another. The feelings were expressed in such a nebulous manner that it was impossible to categorize the statements with enough accuracy to permit tabular presentation. Generally the students seemed to be saying that everyone must do something for the good of the group.

4. Definiteness of personal plans.

Study of the data in Table 17 will indicate the extent to which these subjects in their senior year of high school had made plans which they could begin to carry out within a year. Twenty-five of the subjects omitted all mention of personal plans and wrote about general social problems. Of the 101 others almost half stated very specific plans. Many of the others were waiting to hear from colleges about acceptance and were unable to be more specific. Some had real doubts as the result of their experiences at the Laboratory. One boy expressed this effectively in this way:

Table 17. Definiteness of Plans of Superior Students in Their Senior Year of High School

Classification	Girls N	Boys N	Total N
Had definite plans	29	16	45
Formulated but not definite plans	12	17	29
Planning in nebulous state	14	12	26
Omitted mention of personal plans	9	16	25
Indicated plans would lead to failure	-	1	1
TOTAL	64	62	126

> "It seems as if I am more confused than very about my future vocation, the college I will attend, and what course of study to pursue. But I do not lay the blame for this confusion upon the Guidance Laboratory, except to the extent that I was given

> more information to be confused about. I guess the fact is that I am confused because I am informed about the many opportunities."

The student who indicated that his plans would lead to failure wrote the following paragraph.

> "No matter what I undertake, I will fail at it because I will neither apply myself to it nor take any chances that are given to me to help myself. In short, what I see as my future is failure after failure until I fail at the most important undertaking of a man, life." (This boy withdrew from college after one semester. He was in good academic standing when he returned to his home community to live at home. He was not employed at last resort but had read voluminously.)

5. Interest in education beyond high school.

It is generally expected that superior students will go to college. Since these subjects were not asked to write specifically about their plans for education beyond high school 37 of the subjects did not write about it. (Actually only 7 of the 126 did not continue their education immediately after high school. Four were girls who married within a month of high school graduation.) Of those who did not mention the subject only 2 said that they did not plan to continue and one stated that she was uncertain. The reasons given for continuing education included genuine enthusiasm for study ("I'd like to learn, learn, learn, for the rest of my life."), desire for a comfortable life, preparation for a specific occupation, and the opportunity to talk and know people of different backgrounds and ideas.

6. Concept of vocational roles and range of vocational aspirations.

The data in Table 18 raises some interesting speculations. Thirty-four of the 126 subjects made no reference to their future vocations. Just why they omitted mention of this phase of the future as they saw it is not clear since vocations may play a very important part in one's life and it seems unlikely that these subjects had not considered vocations in their planning. Perhaps they just assumed that, since they had been discussed frequently at the Laboratory, they could omit mention of it in the essay.

Of the 92 subjects who did mention vocational roles only 7 indicated conflict about the personal and social roles. About two-thirds saw that they might contribute to society in the occupations which they chose. These were the subjects who in writing about their vocational choices made statements such as, "My obligation is to contribute to society as much as possible," or "My place in the future is to help others for their future." Only 17 of the 92 who wrote on this matter expressed strictly personal interests "making money," "rising above the masses," and "I want to be looked up to as a businessman of high status." Five of the girls suggested that they considered the preparation for a vocation as insurance ("I want something to fall back on in my marriage if I ever need to.") Only 7 of the 94 indicated that they were not sure what role they might play after they had entered their chosen vocations.

Table 18. Concept of Vocational Roles Reported by Senior High School Superior Students

Classification	Girls N	Boys N	Total N
To satisfy personal needs	8	9	17
To satisfy needs of society	36	27	63
To rely on if need arises	5	-	5
Indecision about vocations	5	2	7
Not mentioned	10	24	34
TOTALS	62	62	126

The vocations given as choices by the subjects included: ambassador, artist, business administrator, engineer, executive secretary, housewife, librarian, mathematician, minister, medical technologist, nurse, physician, psychiatrist, politician, public relations, scientist (various specialties), social worker, teacher (various levels), and United Nations representative.

7. Desire for family membership.

Ninety-one of the 126 subjects made no mention of family membership in the future. Of the 35 who did write about it, 28 were girls. All but one of the 35 indicated that raising a family was an important part of their plans with only one boy stating, and rather positively, that he did not expect to become a family man.

One must wonder why more than half of the girls did not mention marriage and a family in the future as they saw it and their place in that future. Perhaps it was just too obvious and those who did make such statements as, "Naturally, I plan to marry" and "I think the ultimate goal of every girl's life is to marry and have a family" were just stating explicitly what others just assumed. Perhaps the boys became so concerned about their education and vocational roles in the future that marriage seemed rather remote. Or, perhaps they, too, just assumed that readers would recognize without explicit mention that marriage would become part of their future.

8. Unique responses and particular concerns or problems.

Many of the responses were difficult to classify. The following quotation from one essay raises some interesting questions.

> "I want to become a librarian and I hope to do well - but I have no desire to be a dedicated career librarian. In fact I wouldn't care whether I worked at all once I got married. (But I think library work is a career I could go back to) Nevertheless, I don't think college is either a waste of time or money for me if I become 'just a wife and mother.' College will be important to me, not so that I can think of Browning while I wash dishes, but so that I can think, period! I hope to be intelligent, interested, happy and married."

One striking example of maturity in writing about the future that defied the usual classification follows.

> "Until this year the future to me has been tomorrow or the next day, or maybe in ten years....but this year, somehow, my concept has changed. I see my future as a thing of the present. My life right now, what I'm thinking and saying, are all influences upon what I will become...Each new thing or experience for me is part of my future."

Occasionally a student expressed concern about some personal problem such as having a short temper, job security, lack of adequate study habits, inadequate financial support for college or lack of a specific vocational choice. Still others made much of such social problems as the population explosion, the nation's destruction or the national economy. Although they did not appear frequently enough to make tabular presentation desirable there is no question that these were matters of real concern to the writers. It cannot be assumed, however, that those who chose not to write about such concerns were not concerned.

Finally one becomes impressed in reading the essays with the thoughts and ideas of these brilliant students. The 2 quotations which appear below are worth consideration.

> "The future offers the excitement and anticipation arising from the unknown, the uncertain, the insecure. It holds promises of new places, new ideas, and new people; it will measure the progress toward preconceived goals; it will value any attainments reached; and it will record a change in attitudes."

> "Starting college this fall is the beginning of independence and self-sufficiency in my own life. I shall be the one who will be determining the decisions that will affect my future. Moreover, the influence that naturally stems from the home will be in a sense cleaved, and I will resort to new and varied sources for my interests. The future, then, is a means for me to shape my individuality.

Use of the Senior Essay in the Discovery of Potential Social Contributors and Non-contributors

In the analyses described in the previous section it became obvious to the readers that some of the senior student essayists seemed to express genuine desires to contribute to technical or human progress, while others seemed not to be concerned with such matters. Thirty subjects (16 boys and 14 girls) were selected as contributors based upon their statements. Some representative quotations from those classified as contributors follow.

> "When man in general becomes mature enough to take the place in the world intended for him, major troubles will cease. I feel that it is my obligation to do my part to create a condition like this. I must attempt to be a leader toward a purer democracy, voice my opinions on all subjects, and in every way possible try to create a better world. I can do this in many ways. Number one on the list would be to get a good solid education. I should be active in public affairs. I should work to see that competent men get into public office."

> "In the obscurity of the world there is some chance for there are many people who show a power to understand. Understanding might be the solution to co-

> existence. Co-existence is almost the only solution for world peace. I would like to see myself as one of the people who were trying to develop universal understanding."

Having selected the potential contributors, the essays were read again. The first 16 boys and 14 girls who appeared to be non-contributors, for the comparative study of those who expressed interest in aiding others, and those who, while they may have indicated that there was need for improvement of society, did not suggest that they planned to contribute to it. Samples of statements of non-contributors follow.

> "When I complete high school, I will enroll in a college which probably will be the University of Wisconsin. I'm hoping for a large enough scholarship so that I won't have to work except in the summer. I'm not quite sure what courses of study I'm entering but an idea in Chemistry. During the summer months I'm either going to continue working for the Wisconsin Telephone Company or get a job at Yellowstone with a few of my friends. With graduation from college I will work for a few years before getting married. My future still isn't very clear, so it is just one big 'ad lib' until I finally decide."

> "As I see it the next 20 years the world will be much more complex and even the high school graduate will be unable to find a job. Job security will be even hard to find for the college graduate unless he has specialized in some field and has gone on to get his masters or doctors and then it will be relatively simple for him to find job security. He will then have a good chance to broaden his interests and abilities."

In order to check on the original selection of 2 groups of 30 contributors and non-contributors the following procedure was used for the 2 graduate students who were to check the selection process.

1. Contributors and non-contributors' folders were removed from the files plus a random sample of 10 students who were not classified as contributors or non-contributors.

2. The criteria for selection of contributors and non-contributors was explained to the readers.
3. They were asked to select independently those students who they felt were contributors or non-contributors on the basis of the explained criteria.
4. They were also informed that if they felt that some essays could not be classified in either category that they should place them in another category.

Differences between the original selector's selections and those of the 2 additional graduate student readers were remarkably close. The original reader and one graduate student agreed in 96 percent of the cases. Between the original reader and the other student there was perfect agreement.

With the selection of 30 contributors and 30 non-contributors completed the next step required the study of characteristics of each group in terms of a sample of the data available about the subjects in the Laboratory files.

Since this sub-study was an exploratory study in the use of personal documents and since the numbers were too small to encourage generalization, the details of all comparisons made between contributors and non-contributors are not presented in detail here. Instead of reporting the results as findings, it appears better to present them as hypotheses which can be investigated further. The hypotheses given below stem from differences that appeared to be large enough that they seem not likely to occur by chance in larger populations. They suggest that, if studies of larger groups were made by use of null-hypotheses studies, they would result in rejection of such hypotheses. Such studies are planned at the Laboratory as the populations are increased annually. Until such time, however, the following statements of hypotheses are presented for consideration in the hope that others may conduct similar studies. The extent to which senior superior students express intention to contribute to society after completing high school is significantly related to:

1. Terman Concept Mastery Test score. In general the higher the score the greater the expressed intention to contribute to society.
2. Amount of time spent in homework. The students who accepted the responsibility of homework indicated greater intention to contribute.

3. Choice of mathematics as favorite subject and rejection of communications (English, journalism, language) as prepared field of study.
4. Higher grade point averages over the 4 years of high school.
5. Choice of non-fiction over fiction for out-of-school reading. More contributors preferred reading to other leisure-time activities.
6. Choice of careers that are commonly perceived as occupations in which workers are devoted to service to others and improvement of society.
7. Holding of a part-time job while in high school.
8. Preference for music and speech as extra-curricular activities in school.
9. Intention to attend a college or university.
10. Position in family. Eldest students in a family were more likely to be intended contributors.
11. Occupations of fathers. Fewer fathers of intended contributors work in skilled and unskilled occupation.

GENERAL SUMMARY

In reporting such pilot studies as those presented above the writer was fully aware of the limitations of data obtained from personal documents, and was particularly concerned about the possibility that students might try to create a favorable impression or cover up their true feelings. The fact that the students later spent considerable time with counselors who were alert to the possibility of deception and that these counselors had conferences with parents and teachers seemed to reduce the amount of misrepresentation to a minimum. In general the counselors felt that the subjects revealed genuine attitudes as they wrote the essays in their own words and in the length, detail, and direction of their own choosing. Further studies of the subjects at the Laboratory and in follow-up procedures are being used to appraise the validity of the findings noted above. It seems likely that they will be verified.

Details of the findings have been reported in each of the subsections of the chapter and will not be repeated here. They do suggest that analyses of personal documents written by superior students provide evidence about their attitudes that counselors should not ignore. They seem to offer information that cannot be obtained by the highly structured instruments that are commonly employed. And, in one sense, they provide an antidote to the overly-structured approaches to the study of youth that have become popular. They do seem to provide a way to see, in part, the inner half of an adolescent that does not always become apparent to those who observe and test him.

Sample Senior Essay
(not edited)

Name DAVID

You will be given one hour to write on the following topic:

THE FUTURE AS I SEE IT AND MY PLACE IN THAT FUTURE

"I have had for several years a general idea of what I want to do as a career. I have a deep concern for the welfare of the human race and I want to do something to improve it. I also have an immoveable faith in democracy and the American political system as a form of government. Thus,I will probably go into active politics for the purpose of upholding that system and using it to help all men have equal opportunity in a world of peace.

This probably sounds high-flown and idealistic and perhaps even unrealistic. Perhaps I still think that I have to solve all the world's problems myself. But the world's problems will only be solved if there are people who are working to solve them. I have some very strong and basic convictions about the direction along which the solving of these problems should lie, I feel that I have the ability and the imagination, coupled with considerable experience in the practical political science of getting things done, which will allow me to make some useful contribution along these lines.

The two main areas in which I see a place for myself with these ambitions are in active politics and in the foreign service, preferably in Africa. At the present time I lean toward active politics; perhaps after I have lost a couple of bids for elective office I will move into the foreign service as a career.

Between now and the time I embark on a career I intend to get at least four years of a college education and spend a couple of years in the Peace Corps. I do not intend to spend any time in the armed forces unless I have to, for I feel that I can make a greater contribution to the human race by working for peace than being part of a military preventive.

BIBLIOGRAPHY ON PERSONAL DOCUMENTS

1. Allport, Gordon W., *The Use of Personal Documents in Psychological Science*. New York, Social Science Research Council, 230 Park Avenue, 1941.

2. Barbe, W. B., "Study of the Family Background of the Gifted," *Journal of Educational Psychology*. 47:302-309; 1956.

3. Beaglehole, E., "Note on Social Change and Education," *Journal of Educational Sociology*. 29:316-320; 1956.

4. Creegan, R. F., "The Chronological Analysis of Personal Documents," *Journal of Abnormal Social Psychology*. 39:244-266; 1944.

5. Cromett, Harvey J., "An Exploratory Study of Some of the Reading Practices of a Selected Group of Superior Students," *Unpublished Seminar Report*, University of Wisconsin, 1959.

6. Danielson, Paul J., and Rothney, John W. M., "The Student Autobiography: Structured or Unstructured?", *Personnel and Guidance Journal*. 33:30-34; September, 1954.

7. French, J. F., and Steffen, H. J., "Interests of Gifted Adolescents," *Personnel and Guidance Journal*. 38:633-636; April, 1960.

8. Jackson, D. M., and Messick, S., "Content and Style in Personal Assessment," *Psychological Bulletin*, Vol. 55, No. 4; July, 1958.

9. Jessor, R., "Phenomenological Personality Theories and the Data Language of Psychology," *Psychological Review*. 63:173-180; 1956.

10. Jersild, A., *In Search of Self*. New York, Teachers College, Columbia University, 1942.

11. MacCurdy, R. D., "Individualism for the Gifted Child," *Phi Delta Kappan*. 41:408-409; June, 1960.

12. Newland, T. E., "Implications of Research in the Area of the Gifted," *Exceptional Child*. 25:195-198; January, 1959.

13. Rothney, John W., and Roens, Bert A., Counseling the Individual Student, New York: Dryden, 1949.

14. Rothney, John W., Guidance Practices and Results, New York: Harper & Brothers, 1958.

15. Strang, Ruth M., The Adolescent Views Himself. New York: McGraw-Hill, 1957.

16. Strang, Ruth M., "Gifted Adolescents View of Growing Up," Exceptional Children. 23:10-15; October, 1956.

17. Terman, L. M., "The Discovery and Encouragement of Exceptional Talent," American Psychologist. 9:221-230; June, 1954.

CHAPTER III

STUDY OF ORAL PROBLEM-SOLVING PERFORMANCES OF SUPERIOR HIGH SCHOOL STUDENTS

A review of the research on problem-solving revealed few attempts to study processes while the subjects verbalized their activities. Studies involving other than manual manipulation of apparatus or multiple-choice items were rarely reported. No studies in which superior high school students solved problems while thinking aloud and without the use of some apparatus were found. Although few studies relevant to the purposes of this investigation were reported, some of the techniques of scoring and descriptions of behaviors provided suggestions for techniques that might be used in this study. Since none of the investigators confronted their subjects with as great a variety of problems as those used in this investigation, it was necessary to develop new techniques and modify some that had been used.

At the Research and Guidance Laboratory at the University of Wisconsin in 1958, Koplitz reported in an unpublished study that he found differences in problem-solving behavior between ninth grade students classified as superior and average by their teachers. Fifty average students were used as a control group and 50 superior students composed the experimental group. Using problems selected from standardized achievement and intelligence tests, the author hypothesized that there were no differences between the problem-solving performances of superior and average ninth grade students. A check list was employed to facilitate interpretations of the behavior and the problem-solving situation was tape recorded to provide a more detailed and permanent record of thinking out loud. Numbers of right and wrong answers were computed and used to rank the students.

Differences between the mean scores of the average and superior ninth grade students were significant beyond the .001 level. The author indicated that superior students not only achieved at a higher level but also exhibited different behavior than average students. Average students were differentiated from superior students on the following behaviors:

1. Misunderstood directions.
2. Misunderstood specific problems.
3. Failed to apply relevant information to problems.
4. Misunderstood ideas in the problems.
5. Spent less time in thinking through the problems.
6. Avoided use of systems in thinking through problems.

7. Failed to carry reasoning through to correct answer.
8. Lacked self-confidence.
9. Revealed inappropriate attitudes toward problems.

No attempt was made to analyze the differences in behavior further. The investigator concluded that the oral problem-solving technique could be used in identifying and counseling superior high school students.

A review of related research indicated that investigators of problem-solving had employed several techniques. The most frequently expressed emphasis was on the conditions affecting problem-solving. Although it was evident that investigators attempted to gain insights and verify observations by having subjects justify their methods of solving problems, few studies employed "thinking aloud" in order to study processes. The manipulation of apparatus, construction tasks, and multiple-choice problems similar to those found on achievement tests were the tasks employed most commonly in studying problem-solving. Investigators indicated that time, difficulty of the tasks, age of the subjects, attitudes toward the problems and sex differences were variables that needed further research. They failed to agree on the definitions of thinking behavior and conflicting evidence was found when experimenters attempted to duplicate other studies. Methods of improving observations of behaviors and the practices employed to train observers were seldom reported.

Population, Instrument, Procedures

In this study the investigator confronted the superior high school students with situations that required the subjects to give the only correct answer. The subjects thought out loud while they solved the problems which were designed to require no higher level of knowledge than is required in reading and arithmetic. The purposes were to explore the use of the oral problem-solving technique with superior high school students to determine:

1. Sex and grade differences in problem-solving performances.
2. Relationships between performances on the problem-solving exercises and scores on various objectively scored tests.
3. Relationships between procedures used in problem-solving and success in it.
4. Confidence of subjects in their answers to problem-solving exercises.

Population of the Study

Two hundred superior high school students (25 boys and 25 girls from each of the ninth, tenth, eleventh and twelfth grades) were selected randomly from a total of 474 subjects who took the oral problem-solving exercises to participate in this experiment. Since as many as 16 subjects visited the Laboratory in a single day, there were times when the oral problem-solving activity could not be administered to each student. Under these circumstances the subjects were selected randomly. Each subject was given a number that coincided with a marble in a box. An opening at the top of the box permitted only one marble to fall out at a time. If 12 subjects were to participate in the exercise, 12 marbles were removed. The numbers on the marbles identified the subjects who were to participate. Since visits to the Laboratory were scheduled prior to, and without consideration for, the design of this study, it seemed unlikely that selection bias would affect the results.

Methods employed by the schools to select students who visited the Laboratory were not standardized. As previously described, the schools used their own methods of selection. Laboratory recommendations and reports from the schools indicated that the students were usually selected on the basis of the results of standardized tests, performances in school, and teachers' recommendations. The mean scores of the subjects on the total score section of the SCAT 2A indicated that they generally ranked above the 95th percentile.

The Oral Problems

The instrument employed in this investigation consisted of 12 situations. They were selected because they required the subjects to apply general problem-solving skills rather than knowledge of specific course work. Only one correct answer could be given for each of the problems. It seemed that if subjectswere required to reach an answer employing only the information given the resulting data would be representative of problem-solving skills.

Enough information was given to solve each problem. The knowledge required to reach correct answers to each seemed likely to be common to high school students. Simple arithmetic, family relationships and the right angle corner of a square appeared to represent the most difficult concepts involved in solving the problems. As in any problem in which directions are written, elementary reading skills were required. It was unlikely that subjects who performed at a superior level on standardized tests would lack sufficient reading skill to understand the assigned tasks.

Fifteen problems were selected from magazines, intelligence tests, newspapers, puzzle books and mathematics publications. The 15 problems were administered experimentally to 82 superior students selected from the Laboratory population. This trial study resulted in the elimination of 3 problems because they were apparently trick questions. The remaining 12 were then arranged in order of increasing difficulty for presentation to the subjects of the experiment.

Problems

The first problem, borrowed from a well-known test, was designated as the robbery problem.*

> Tom's house was burglarized last Saturday. Tom was at home all of the morning but out during the afternoon until 5 o'clock. His father left the house at 3 o'clock and his brother was there until 4. At what time did the burglary take place?

The second problem in order of difficulty was called the check problem. It seemed to require the establishment of simple relationships, but the information was arranged in a more complex fashion than in the first problem.

> One day last week a man went to town with only five dollars in his pocket, but returned with fifteen. He bought a pair of shoes and some meat in the market. He also had his eyes examined. It so happens that the man gets paid every Thursday by check, and the banks in this town are open to cash checks on Tuesday, Friday and Saturday only. The eye doctor is not in his office on Saturday and there is no market on Thursday or Friday. What day did the man go to town?

The next problem, third in order of increasing difficulty, confronted the subject with a sudden change in method of presentation. The first two problems were lengthy and required the subjects to relate a number of facts. The third problem, grandfather, was brief but required the subject to define four family relationships as they pertained to one individual.

> What relation to an individual is his sister's grandfather's only child's wife?

---

*The underlined word in each of the paragraphs in which the problem is described will be used henceforth as a short method of referring to problems.

The fourth problem (ducks) confronted the subjects with an apparent contradiction. The subjects were to give an answer that indicated there was not a contradiction in the problem.

> Two fathers and two sons went hunting. Each shot a duck and none shot the same duck. However, only three ducks were shot. How could this happen?

The fifth problem (red cows) was the only example of a presentation of alternate choices, one of which was the right answer.

> Four white cows and three red cows gave as much milk in five days as three white cows and four red cows gave in four days. Do white or red cows give the most milk?

The sixth problem, siblings, was one of three family relationships problems.

> In the Robinson family, each daughter has the same number of brothers as she has sisters, and each son has two more sisters than he has brothers. What is the least number of children in the family?

The seventh problem administered to the subjects was the truck problem.

> A truck starts out for the city dump, 10 miles away, going at a speed of 5 miles an hour. At the same time, a horsefly that can fly 20 miles an hour leaves the truck and flies to the dump, immediately turns around and returns to the truck, and keeps shuttling back and forth this way until the truck arrives at the dump. How far does the fly travel altogether?

The eighth problem dealing with family relationships was the in-laws problem. This problem was selected because it confronted the subjects with an unusual family relationship. It was assumed that each of the superior subjects was familiar with brother-in-law and stepmother relationships. The subjects were to interpret family relationships in an unusual situation. This family relationship problem seemed to be in a category separate from numbers 3 (grandfather) and 6 (siblings).

> A wedding produced the following results. The father of a young woman became her husband's brother-in-law, and her husband's sister became her stepmother. How could this happen?

The ninth problem administered to the subjects was the socks problem.

> There are six white socks and six black socks in a drawer. If one reaches into the drawer in the dark, so that he is unable to see them, how many socks must be removed to make sure he has a pair of matching socks?

The next situation was the moat problem which was the tenth in order of administration. The description of the situation seemed to offer a humorous touch to the problem-solving experience. Diversity of problems and their descriptions seemed important when working with superior students.

> A prince is being pursued by robbers. Upon reaching the castle, he finds the drawbridge up and no one attending it. The castle is surrounded by a square moat three feet from the castle and 8 feet in width. It is filled with dangerous reptiles. The prince finds two boards each of which is seven feet eight inches long. How can he cross the moat?

The railroad problem, number 11, required the subject to keep facts in order and to understand the location of points in the problem, or use the facts and arrive at an answer by relating one fact to another.

> A railroad line runs from Pine to Oak. Elm is on this line just half the way between Pine and Oak. Maple is just as far from Pine as it is from Elm and Elm is as far from Maple as it is from Oak. If it is twenty miles from Pine to Maple, how far is Maple from Oak.

The final problem administered to the subjects was number 12, cows in line.

> When driving some cows, a boy was asked how many cows he had. He replied, "When they are in line

> there are two cows ahead of a cow, two cows behind a cow, and one cow in the middle." What is the least number he could have?

The description of the problems indicates that each of the problems confronted the subjects with different situations. It appeared that each problem presented them with alternate means of attacking it. If the subjects followed the directions for the exercises there was only one correct answer for each problem.

A pilot study of superior students was used to determine the difficulty of the problems and a time limit of 3 minutes was established for each. The percentages of subjects who answered each of the 12 problems correctly when they were administered in the indicated order are reported in Table 1. The percentages of correct responses indicate that only the eighth problem, in-laws, the fourth problem, ducks, and the twelfth problem, cows in line, changed more than one rank in order of difficulty. The order in which they were presented seemed to be similar to the difficulty levels for the subjects in this investigation.

Table 1. Percentages of Two Hundred Students in Grades Nine through Twelve who Answered the Exercises Correctly

Problems	Percent Correct	Difficulty Rank
1. Robbery	88	12
2. Check	74	11
3. Grandfather	68	10
4. Ducks	41	4
5. Red cows	57	8
6. Siblings	44	6.5
7. Truck	42	5
8. In-laws	58	9
9. Socks	44	6.5
10. Moat	24	2
11. Railroad	14	1
12. Cows in line	35	3

Confidence Level

After the subjects gave an answer to each problem the administrator asked the subjects about their certainty of the correctness of their answer. A 5 x 8 card was presented to the subjects after each problem had been given so that they could choose from the 3 responses that are listed below.

A. The answer I gave is the only possible answer.
B. My answer seems right but if I had a chance to do it over again I might find that it was not correct.
C. I am sure that my answer was not correct.

Observations During Problem-Solving

The oral problem-solving exercises were administered to each of these subjects individually. The administrators observed the behaviors of the students and recorded their observations on a check list designed for this purpose. Seventy of the 200 subjects were watched by the administrator of the exercises and an independent observer. A one-way vision screen separated the observer from the subject and an administrator. The student's answers were tape recorded and the observer was able to hear the subject's responses through a speaker connected to the recorder. The subjects were informed that tape recordings and observations were being made. These techniques were not new experiences since most of the Laboratory students were frequently observed and tape recordings of interviews were often made.

Observations were made whenever time and availability of staff members permitted. A minimum of 17 observations for each grade and at each grade level was made. The first 70 observations were used in this study. Each of the 8 graduate assistants, who participated in the administration and observation of the problem-solving exercises, administered and observed at least 3 trial performances of one subject. During the 2-week training period, 2 or more assistants observed the problem-solving activity and recorded their observations independently. The protocols were compared and the tape recordings were reviewed as a means of helping the observers gain an understanding of the behaviors exhibited by the subjects, and to gain skill in recording their observations.

During the trial period administrators were observed for the purpose of standardizing methods of administration and observation. Although the formal period of training ended when the first subject in this study was administered the oral problem-solving activity, the investigator observed each administrator at least once, and some as often as 10 times. Observers and administrators usually compared their observations after each session.

Tape recordings were made of each of the problem-solving activities. The first 25 subjects who gave right and the first 25 subjects who gave wrong answers to the red cows problem were selected for tape script analysis. The same procedure was followed for the socks and the railroad problems. By using this means of sampling the selection depended upon the answers to the individual problems. Answers that were given to the socks problem had no effect upon the selection of tape scripts for the red cows or railroad problems.

Subjects' comments about the oral problem-solving activity seemed to offer a means of checking the results of the investigation and, at the same time, provided suggestions that could be used to improve the instrument. At least one of the suggestions offered by a trial subject was incorporated in this investigation. The first question asked the students after completion of the activity concerned their use of mental images in solving the problems.

The inclusion of the question,"Did you ever get a mental picture of the problem situation?", was a result of suggestions made by 4 students in the pilot study. They commented that mental pictures were frequently used, but difficult to describe.

Since the suggestion of mental images arose following queries about techniques the subjects used in problem-solving, the second question was included in this experiment, "Was there anything else that you did or thought during the problem-solving that you haven't told me about?".

The third question was constructed as an open-ended item to enable the subjects to comment in any way they desired about their experiences. Since Laboratory students are frequently asked to comment on educational procedures employed, it seemed that they would express themselves freely on this item. If a large percentage rejected this technique of exploring their problem-solving performances it would suggest that little value could be obtained from continuing the experiment. If subjects failed to verbalize their thought processes little could be gained from the experiment.

Procedures

The oral problem-solving activity was scheduled throughout the day. The time of administration depended upon the availability of graduate assistants and the subjects' schedules of activities at the Laboratory. Subjects were assigned to administrators of the exercises by chance. The subjects were listed alphabetically on the schedule board across from the names of the administrators. It seemed that this method of assignment was not likely to produce assignment biases.

A standardized set of directions was developed through experimental administration, and revised on the experiences gained in the pilot study.

In order to obtain the fullest cooperation of the subjects it seemed that they should be informed of the purpose of the exercise. It also seemed expedient to prepare the subjects for the exercise by outlining the conditions under which the problems would be solved. The directions were read to the students and they were asked if there were any questions. If the subjects asked a question, the administrator indicated that he could only repeat the directions that related to the question asked. Subjects accepted this as a requirement of the exercise.

After the directions were read and subject's questions answered, the administrator gave the first problem to the subject. They were each typed on separate 5 x 8 cards. The subject read the problem aloud and immediately started to solve it. A stop watch or timer clock was started as soon as the subject had read the problem to see that he did not exceed the 3 minute time limit. If the subject failed to give an answer or did not give up before time expired, he was notified that time ran out. Subjects were allowed to continue solving the problem if they desired to do so. Processes employed or answers given were not recorded after time expired. Subjects who gave an answer to a problem were then asked to indicate their confidence in the correctness of their response. After the subjects stated their level of confidence they were asked to review the problem. The review was not analyzed in this investigation. The subjects were asked to review each problem as a means of equalizing the opportunities to learn from previous errors or to reinforce their commitment to the methods employed. After the 12 problems had been completed and the subjects had selected their final confidence level and reviewed the last problem, the administrators asked the subjects the 3 questions previously described. This terminated the activity.

Data were collected by the use of 3 techniques. Administrators recorded subjects' behaviors on each problem, right and wrong answers, time taken for each problem, estimates of levels of confidence for each, and answers to the questions at the termination of the exercises for 200 subjects. Each observer recorded the subjects' behaviors, right and wrong answers, and estimates of levels of confidence for each problem during 70 of the 200 problem-solving activities. Tape recordings were made of each activity and 2 independent raters analyzed the tape scripts of a sample of the subjects.

Since there was only one correct answer for each problem the recording of right and wrong answers seemed to present little difficulty. The subjects were not informed of the correctness of their answers until their final test interpretation interview, the last activity of the day. It was important for standardization that the administrators use symbols for answers that the subjects could not interpret in terms of marks for right or wrong answers.

A time limit of three minutes was set for each problem. Experience gained in trial administrations indicated that some subjects would take as much time as they were permitted. It was also discovered that over 95 percent of the correct answers to each problem were given in less than 3 minutes. Consistent time limits were set for each problem so that the subjects could become accustomed to the length of time allowed and to prevent the subject from worrying about the time. Time was recorded to the nearest whole second.

Confidence estimates were recorded by letters. "A" indicated the subjects were certain of the correctness of their answer. "B" indicated they were doubtful. "C" indicated they were sure their answer was not correct. Confidence estimates were obtained for each of the problems.

Seventeen behaviors were included on the observation protocol for each problem. The behaviors were listed in order of their usual occurrence during the problem-solving process in order to facilitate the process of checking the behaviors that were observed. A list of definitions and examples were prepared for each of the behaviors.

Subjects were not permitted to use pencil and paper because it seemed that the use of writing materials would detract from the degree to which the subjects would verbalize their thinking. It was found through trial administrations that subjects tended to use their fingers to count, to indicate locations, or to draw diagrams of the problem situation. If subjects who used these techniques gave more right or wrong answers it might indicate that some form of physical manipulation of figures or locations was an aid or hindrance to problem-solving.

Subjects' reactions to the activity were likely to have some effect upon their performances. Since evaluation of subjects' reactions is difficult to categorize, only 2 such behaviors were entered on the observation sheet. It appeared that observers could determine whether or not subjects enjoyed the challenge of the experience and if they were hesitant or timid during the activity. Those behaviors were not considered as either-or situations. It seemed that subjects could be hesitant or timid while experiencing the new situation of thinking out loud and still enjoy the challenge of the activity. If subjects who exhibited either of these behaviors gave more right or wrong answers it would seem that reactions to the activity would be related to success or failure in it.

After the data on the 200 subjects were collected, the information was tabulated and arranged for statistical analysis. The presentation and discussion of the results are reported in the following pages.

Findings

Percentage of Correct Responses

Subjects responded to the exercises by giving a correct answer, giving an incorrect answer, giving up, indicating that they had heard the problem before, or attempting to solve the problem until time expired.

Subjects gave up in 276 of a possible 2400 (11.5 percent) attempts to solve the problems. Several interpretations of their behaviors were given by the students when giving up before time expired. Since no provisions had been made for recording the interpretations further analysis of the data was not attempted.

Every problem except robbery and railroad had been heard by at least one student. Problems check, grandfather, red cows, siblings, truck, and in-laws were heard by less than 5 subjects. Problems ducks, socks, and cows in line had been heard by 25 subjects or less. Twenty-eight percent of the subjects had heard at least one of the exercises before, and none of the subjects had heard more than two. Subjects were not asked to solve problems that they had heard before.

The percentage of subjects who answered the exercises correctly are reported by grade level in Table 2 .* Tests of significance between percentages were applied to the data. Differences between the performances of students on problems siblings, in-laws, and cows in line, indicated that seniors performed at a significantly higher level than sophomores on the sibling problem; a significantly smaller percentage of freshmen answered the in-laws problem correctly than did seniors, and a significantly larger percentage of seniors and sophomores answered correctly the twelfth problem, cows in line, than did freshmen. When the data were distributed by sex significant differences beyond the 5 percent level were found between the performances of the sexes on the same 3 problems. It seemed that the significant grade level differences might be due in part to the differences in performances between the sexes.

---

*The problems in the following tables are listed in order of increasing difficulty.

Table 2. Percentages of Fifty Students in each of the Grades Nine Through Twelve who Answered the Exercises Correctly

Exercises	Ninth Grade	Tenth Grade	Eleventh Grade	Twelfth Grade
Robbery	86	88	84	94
Check	74	72	72	78
Grandfather	76	66	72	78
In-laws	52	56	56	68
Red cows	50	54	56	68
Siblings	48	28	42	56
Socks	40	36	48	44
Truck	48	36	48	38
Cows in line	22	40	36	40
Moat	14	14	20	46*
Railroad	14	18	20	06+

+Significance was not computed when frequencies were less than five.
*Indicates a percentage significantly different beyond the 5 percent level from each of the other 3 grade-level percentages for the same problem.

The results of the comparisons presented above indicated that significant differences, beyond the 5 percent level, between the performances of subjects in one grade and the other 3 grade levels were found for only problem 10, moat. This problem was near average difficulty for the seniors while it was one of the 2 most difficult problems for the other classes.

It appeared that grade level was not related to success on the 11 other problems. This finding suggests that length of school attendance has minor effects upon the skills involved in solving these problems. One might conjecture that the superior students had acquired sufficient skills in problem-solving by the time they were in ninth grade to enable them to utilize newly acquired knowledge in related problem situations, such as achievement tests, compositions and class demonstrations. Perhaps grade level differences would be apparent in problems dependent upon knowledge of subject matter, rather than the exercises used in this experiment.

Since the problems were selected on the basis of being dissimilar it might be assumed that one or more of the problems included information or were so structured that a sex difference might result. Percentages and differences between percentages for the sexes are reported in Table 3. The differences between the percentages for the sexes were significant beyond the 5 percent level on problems 4, ducks, 6, siblings, 8, in-laws, and 12, cows in line. Problems 6, siblings, and 8, in-laws, dealt with family relationships and the girls performed at a significantly higher level than the boys. Girls seemed to feel more at ease in working with family relationships than did boys.

Table 3. Percentages and Differences Between Percentages of One Hundred Boys and One Hundred Girls who Answered the Twelve Exercises Correctly

Exercises	Boys	Girls	Differences
Robbery	87	89	- 2
Check	73	75	- 2
Grandfather	70	65	+ 5
In-laws	48	68	-20*
Red cows	63	51	+12
Siblings	36	51	-15*
Socks	43	45	- 2
Truck	41	44	- 3
Duck	49	33	+16*
Cows in line	48	22	+26*
Moat	28	20	+ 8
Railroad	13	16	- 3

*Significant beyond the 5 percent level.

Problems 4, ducks, and 12, cows in line, were answered correctly by a significantly larger percentage of boys than girls. Analysés of tape scripts and protocols offered little help in interpreting these results. The mean times for correct answers varied from year to year for both sexes. The cows in line problem was one of the 2 exercises answered most quickly by boys and girls in each grade, and the ducks problem was of about mean time difficulty for both sexes.

Time Required for Correct Responses

The time required for the subjects to give correct answers was recorded. A limit of 3 minutes was established for each of the problems because most of the trial subjects who arrived at correct answers did so within this limit. Mean times in seconds and the differences between means of boys and girls in ninth and tenth grades are reported in Table 4. A test of significance was applied to the differences between means of the sexes for each grade level. No differences significant beyond the 5 percent level were found. This finding indicates that time differences were not closely related to the sex differences found previously. It also suggests that time was not as important as might be anticipated for such oral problem-solving exercises.

Table 4. Mean Times in Seconds and Differences Between the Means of Ninth and Tenth Grade Boys and Girls who Answered the Exercises Correctly

	Grade 9					Grade 10				
	Boys		Girls		Diff.	Boys		Girls		Diff.
	N	Secs.	N	Secs.	Secs.	N	Secs.	N	Secs.	Secs.
Robbery	22	34	21	26	+ 8	21	32	23	29	+ 3
Check	19	60	18	45	+15	18	63	18	49	+14
Grandfather	18	41	15	30	+11	18	45	15	35	+10
In-laws	10	83	16	83	0	11	37	17	58	-21
Red cows	16	47	9	39	+ 8	13	50	14	46	+ 4
Siblings	11	40	13	63	-23	6	84	8	40	+44
Socks	8	40	16	20	+20	11	27	7	14	+13
Truck	10	51	14	78	-27	11	88	7	48	+40
Ducks	10	56	8	36	+20	13	54	8	51	+ 3
Cows in line	11	18	1	31	-13	13	25	7	17	+ 8
Moat	5	54	2	48	+ 6	4	98	3	74	+24
Railroad	4	106	3	76	+30	4	102	5	109	- 7

Another interesting comparison was that the differences in confidence levels between boys and girls remained relatively stable for the 4 years. Inspection of the data indicated that there were no consistent differences between the mean times for boys and girls over the 4 years that approximated the magnitude needed for significance beyond the 5 percent level. Interpretation of the findings was further limited by the directions for the administration of the exercises. Subjects were instructed that time was important, but getting the correct answer was the most important task. They were also reminded throughout the exercise that thinking out loud was very important. Time was called when the 3 minutes had elapsed, but the subjects were allowed to pursue an answer if they so desired. A record was not kept of those who gave correct responses after time had expired. The design of the study and the directions for the exercises seemed to relegate time to a position of minor importance to the student when compared to thinking out loud and arriving at a correct answer. Variability in time might be a result of the subject's interpretations of the importance of time in the problem-solving exercises.

Subjects' Confidence in the Correctness of Their Answers

Three statements described previously were designed to determine the subjects' confidence in their responses. Numbers, percentages and differences between percentages of boys and girls who gave right or wrong answers and who also indicated they were certain of the correctness of their answers to the exercises are listed in Table 5. The numbers in Table 5 refer only to those students who gave an answer that could be recorded as right or wrong. It did not include subjects who gave up, who had heard the exercise before, or who did not solve the problem before the time expired.

Table 5. Numbers, Percentages and Differences Between Percentages of Boys and Girls who Gave a Right or Wrong Answer to the Problems and who Indicated They were Certain of the Correctness of Their Answers

Problems	Boys N	Boys %	Girls N	Girls %	Diff. %
Robbery	83	84	75	75	+ 5
Check	50	50	42	43	+ 7
Grandfather	63	64	49	50	+14*
In-laws	40	56	37	45	+11
Red cows	37	38	19	21	+17*
Siblings	46	52	36	39	+13
Socks	64	77	50	57	+20*
Truck	27	39	16	23	+16*
Ducks	55	71	39	58	+13
Cows in line	77	80	55	58	+22*
Moat	26	35	13	19	+16*
Railroad	21	28	13	18	+10

*Significant beyond the 5 percent level.

Tests of significance were applied to the differences between percentages of boys and girls. The boys apparently selected A, to indicate they were certain of their answers, more frequently than did the girls for every problem. Significantly more boys than girls indicated that they were certain of their answers for problems 3, grandfather, 5, red cows, 9, socks, 7, truck, 10, moat, and 12, cows in line. Four of these 6 exercises were among the most difficult problems. When the second choices (doubtful of the correctness of their answer) were tabulated for comparison of boys and girls, the girls seemed to choose this category more frequently than did the boys for every problem. The results of these tabulations are recorded in Table 6. Significant differences between the percentages were found on problems 1, robbery, 5, red cows, 6, siblings, 9, socks, and cows in line. Only 2 of these significant differences appeared among the 6 most difficult problems. The difficulty of the problems seemed to have little effect upon the levels of confidence expressed by either the boys or the girls.

Table 6. Numbers, Percentages and Differences Between Percentages of Boys and Girls who Gave a Right or Wrong Answer to the Problems and who Indicated They were Doubtful of the Correctness of Their Answers

Problems	Boys		Girls		Diff.
	N	%	N	%	%
Robbery	13	13	25	25	-12*
Check	46	47	53	55	- 8
Grandfather	31	32	45	45	-13
In-laws	27	38	39	48	-10
Red cows	54	56	68	74	-18*
Siblings	32	36	47	51	-15*
Socks	17	20	31	36	-16*
Truck	35	51	43	61	-10
Ducks	16	21	20	30	- 9
Cows in line	18	19	38	40	-21*
Moat	42	56	33	48	- 8
Railroad	44	59	50	69	-10

The numbers and percentages of boys and girls who selected C to indicate that they were sure their answer was not correct indicated that there was little difference between the percentages of boys and girls who chose this response. Significant differences between boys and girls were found on the moat problem. Boys were certain their answers were wrong more than the girls. No differences were found to be significant beyond the 5 percent level for the remaining 11 problems.

Subjects' Comments About the Problem-Solving Activity

After completing the exercises the subjects were asked if they had mental images of the problem situations. Since the responses to this question were similar for boys and girls they were combined. Percentages of the 200 students who responded in the manner indicated to the question, "Did you ever get a mental picture of the problem situation?" were recorded in Table 7. The category "most of the time" included students who mentioned, most or all of the time, or enumerated 7 or more problems in which they pictured mentally the people or objects presented in the exercise. The category "some of the time" includes responses which indicated that they had mental pictures of the problem situation in 6 or less problems or indicated that they had mental pictures of some of the problem situations without enumerating them. Some examples of explanations for the answers given to the question about mental images are included below.

Table 7. Percentages of Two Hundred Subjects who Responded in the Manner Indicated to the Question, "Did You ever get a Mental Picture of the Problem Situation?"

Responses	N	Percent
Most of the time	88	44
Some of the time	82	41
Attempted to get mental pictures	8	4
Mental pictures were necessary for me to solve the problems	4	2
Mental pictures were not of objects in the problems but I used numbers or hypothetical images	4	2
I had mental pictures of the problems I answered incorrectly	2	1
Never had mental pictures	8	4
No comment recorded	2	1
Question was not asked	2	1

> "I could see the supermarket and the bank in our town when I read the problem." (check problem)
>
> "All I could see were five cows in a cluster, you know two ahead, one in the middle and two, side by side, in the rear." (cows in line problem)
>
> "I saw a flock of ducks fly over and four people shoot and then tried to figure out how all of them could shoot a duck and still only three ducks were shot." (ducks problem)

Only 4 percent of the subjects indicated that they had never had a mental image. Eighty-five percent responded that they had mental images for some or most of the problems. When the other items that suggested students had some mental images were combined, the total of those having had at least some mental pictures was 94 percent. Of the remaining 6 percent, 4 percent had no mental images, 1 percent were not asked and 1 percent did not comment. These results suggested that most of the subjects visualized a picture or an image of objects, people, animals or places, depending on the description of the problem, which they used as a means of solving it. Although mental images were difficult to describe, one of the students explained that they were just like looking into a slide projector. Some indicated that the pictures were in technicolor and others suggested that they could see the images in any color they wanted. A senior boy suggested that drawing mental pictures and day-dreaming were similar processes.

The subjects were asked if they had any comments to make about the exercises. The results were categorized and are reported in Table 8. Thirty percent of the students indicated that problems of this type could be done easier with pencil and paper. Twenty-five percent thought the exercises were fun, interesting or challenging. Nineteen percent replied that this type of exercise made them think but 8 percent indicated that the exercises were of little value. They suggested that other tests measured the same thing and that regular test procedures prepared them better for college entrance examinations or other tests they would have to take in the future.

Table 8. Percentages of Two Hundred Subjects who Responded in the Manner Indicated to the Question, "Do You have any Comments to Make About an Exercise of this Type?"

Responses	N	Percent
If I could have used a pencil it would have been easier	60	30
It was fun, interesting or challenging	50	25
It made you think	38	19
I had trouble thinking and talking out loud	25	12
It was just like all the other tests	15	8
No comment	10	5
Question not asked	2	1

Only 6 percent of the subjects indicated that thinking out loud hindered them in solving the problems. Some unique and interesting statements were hidden through categorization. The following replies were among the unique responses.

"My thoughts are private, I don't like people staring at me when I'm working."

"I never had any preparation for this type of test."

"There were so many pressures, how can a person think?"

"It was the best challenge I've ever had."

The majority of the subjects were not disturbed by the exercise. Although some of the students indicated that they preferred regular objectively scored tests, they seemed to adapt readily to the new situation. The small percentage of students who indicated that they were hindered by thinking out loud suggested that the procedure might be an effective device for studying methods used by superior students in problem-solving. The opinions given by the students indicated the rapport that was developed between them and the administrators of the exercises at the Laboratory.

Oral Problem-Solving and Performances on Standardized Tests

The correlation coefficients between total scores on the oral problem-solving exercise and objectively scored tests were: .193 with total scores on the SCAT 2A, .201 with the D.A.T. Abstract scores, and .249 with scores on the Terman Concept Mastery test. Since the exercises were selected on the basis of apparent dissimilarity and since the odd-even technique had indicated that the over-all reliability was low, bi-serial correlations were computed for each of the exercises that were answered correctly by less than two-thirds of the subjects. Inspection of the exercises answered by more than two-thirds of the subjects indicated that the items lacked sufficient power of discrimination to be of value in bi-serial correlations. Total scores on the Terman Concept Mastery, Watson-Glaser Critical Thinking Appraisal, the SCAT 2A, and D.A.T. Abstract test were used as measures for bi-serial correlation with each of the 12 exercises separately. Total scores were collected for seniors on the Terman Concept Mastery, for juniors on the Watson-Glaser Critical Thinking Appraisal, for sophomores, juniors and seniors on the SCAT 2A, and for the 4 grade levels on the D.A.T. Abstract test. These tests were selected because they seemed to offer a distribution of the types of problems found in test batteries. It will be recalled that the authors of the SCAT 2A indicated that total scores could be used as a measure of total school achievement, the D.A.T. Abstract Reasoning test was supposed to have predictive validity for achievement in the scientific areas and the Terman Concept Mastery test was said to be useful in measuring scholastic aptitude.

Ten bi-serial coefficients were significant beyond the 5 percent level for the 4 least difficult problems and 8 coefficients were significant for the 4 most difficult problems. There seemed to be little relationship between problem difficulty and the number of significant coefficients. The D.A.T. Abstract and the SCAT 2A scores were significantly correlated with 5 of 9 exercises. When the problems treated separately by sex were examined, it was found that girls' performances on the *siblings* and *in-laws* problems correlated significantly with the Terman Concept Mastery, the SCAT 2A, and the D.A.T. Abstract. No important coefficients appeared between the boys' performances on these problems and on the standardized tests. The sex differences previously found with these exercises seemed to be substantiated by these findings.

The significant coefficients between girls' performances on the *siblings* and *cows in line* exercises and the Terman Concept Mastery, the D.A.T. Abstract and the SCAT 2A ranged from .354 to .574. Boys' performances on the Watson-Glaser correlated significantly with their performances on the *ducks* and *cows in line* exercises. The coefficients ranged from .545 to .602. Performances on the *red cow*, *moat* and *railraod* problems were significantly correlated with only one of the standardized tests. The *red cows* exercise was significantly correlated with the D.A.T. Abstract (.324). The *moat* problem was significantly correlated with the SCAT 2A (.327) and the *railroad* exercise was significantly correlated with the Watson-Glaser at the .327 level. Total scores on the Watson-Glaser, SCAT 2A and the D.A.T. Abstract were significantly correlated with the *socks* problem. The coefficients were .376, .201 and .285. Although chances of accurate prediction of success in oral problem-solving from test scores are not high, they do suggest that there is not sufficient independent data obtained from only the right or wrong responses to the oral problem-solving exercises that could not be obtained from administration of the 4 standardized tests; but the significant differences found between the performances of the sexes suggests that oral problem-solving may offer clues about techniques that were not demonstrated by consideration of only the correctness of the answers given.

Observation of Students While They Solved Problems

Administrators of the problem-solving activity recorded the behaviors exhibited by the subjects while attempting to solve each of the 12 problems. It will be recalled that 70 subjects were also observed by graduate assistants while others administered the problems. Each of the workers participated in a training period of 2 weeks and engaged in periodic reviews as checks on methods of administration and observation. It seemed that trained, independent observers could agree on the exhibited behaviors of subjects during the problem-solving activity.

Examination of the data indicated that independent observers and administrators agreed in more than 75 percent of the cases on each behavior. The percentages of agreement ranged from 76 to 100. The greatest amount of disagreement occurred on the "if-then approach" (range of 76 to 91 percent) and "enjoyed the challenge" (range of 80 to 94 percent). Although perfect agreement was the goal, the average agreement of 96 percent suggests that observers could agree on whether or not subjects exhibited certain behaviors during the problem-solving process. There is also the suggestion that thinking out loud may represent an effective means of studying the methods employed by superior students in solving problems.

Disagreement among observers seemed to be closely related to 2 conditions. The first of these was the physical situation surrounding the activity. Observers had different views of subjects and it appeared that one observer could see the subject completely while the other observer could not. Besides location in relation to the 2 observers, there was also a desk between the administrator and the subject. At times the desk seemed to block the vision of one of the observers.

A second factor related to disagreement was the procedure involved in administrating and observing the exercise. It was felt that administrators should attempt to keep the relationship between the subjects and themselves as close to the counseling relationship as possible. If this type of relationship could be approximated, it was felt that the subjects would make their best effort and thereby facilitate the observations of the procedures used in solving the problems. The degree to which the subjects verbalized suggested that the relationships developed were effective in motivating the subjects to think out loud. An interesting observation was made by a visitor who observed the exercises. He indicated that the administrator was so absorbed with the subject that the roof could cave in and he (the administrator) would not notice it.

Inconsistencies in observations seemed to be related to differences in the observers' training. When 8 people were involved at different times in the task of observing subjects' performances on 12 exercises for 40 minutes, there seemed to be many different interpretations for each behavior. The results of the training procedures used in this study suggest that the number of observers should be limited. An optimum number seemed to be 3 to 5, since they would have an opportunity to work with each of the other observers more frequently. Three to 5 observers also provided enough divergence in interpretations. It seemed that even after working with the problems for a long period of time, some facets related to using the given information were not discovered. Familiarity with experimental procedures seemed to be of utmost importance for observer agreement.

Observers apparently experienced little difficulty in noting the behavior "enjoyed the challenge," and observer agreement fell below the 80 percent level in only one problem. Relationships between the subjects and the administrators of the problems seemed to be important. The results of the analysis of agreement between observers suggests that observers can agree above the 70 percent level if the problems are of sufficient complexity and the observers are trained and completely familiar with the problems.

Analyses of Tape Scripts

Analyses of tape scripts were employed to identify the procedures used by the subjects. It will be recalled that 2 graduate students read the tape scripts independently and recorded the behaviors. Each was familiar with the problems and had participated in the administration and observation of the activity. Trial tape scripts were analyzed as a means of training the raters and of studying the behaviors on the check sheet to aid in recording their analyses. Revisions of the check list consisted primarily of re-wording descriptions of behavior and refining the methods of recording them. It seemed that correct answers were related not only to use or neglect of a behavior, but also the degree of use. As an example, 50 subjects may have been observed making interpretations necessary to solve the problem but 25 of these subjects gave right and 25 gave wrong answers. It appeared that subjects who gave right answers made all the interpretations necessary to reach a right answer and those who gave wrong answers made only some of the necessary interpretations. The 3 categories of present, partly present, and absent were used for 10 of the 28 behaviors to indicate whether the behaviors were applied to all the information, part of the information or lack of use of the behavior. Descriptions of behavior such as "misinterpreted information," "emphasized key words in the statement of the problem," "emphasized key words during the solution of the problem," "stated necessary relationships," "stated essential interpretations," "attempted mathematical approach," "made false interpretations," and "checked throughout the problem" seemed to be amenable to rating by degrees.

Raters agreed that the remaining 18 behaviors, including such items as "used mathematical approach," should be reported in only 2 categories, present or absent, in the tape script analysis. The raters reviewed the behaviors identified for every fifth subject to discuss any discrepancies that arose. Differences between the raters and ommissions or mistaken interpretations were discussed. Neither the interpretations made by the raters nor the list of aids for interpretations were revised after the rating of sample tape scripts began.

The results of the analyses of tape scripts is presented and discussed in 2 parts. In order to facilitate the presentation of the results, only those behaviors that were employed by 5 or more subjects who gave right or wrong answers were used. It seemed that frequencies which accounted for less than 20 percent of the subjects would be of little value in investigating the behavior of this sample, but it does not infer that the behaviors omitted were not essential to the individual subjects who employed them. The analyses were undertaken to investigate the behaviors which seemed to be related generally to right and wrong answers. The numbers and percentages of agreement between raters who employed the behaviors and the percentages of agreement between raters who identified them are presented in the first part of this discussion. Since these data appear to be more relevant to numbers of subjects who employed the behaviors and the degree to which the raters agreed upon the identification of them, subjects who gave right and wrong answers and employed the behaviors will be discussed at the same time.

Three tape recordings for 3 of the problems, an easy, an average, and a difficult problem were analyzed. The numbers and percentages of subjects who employed the indicated behaviors and the agreement between raters on the identification of them were studied. Only 9 behaviors were employed by 20 percent or more subjects who gave right and wrong answers. Rater agreement ranged from 47 to 100 percent. Raters indicated that 56 percent of the subjects who gave right answers eliminated alternatives by generalization and 67 percent eliminated alternatives using the step-by-step process. Agreements on these 2 behaviors apparently occurred because of the difficulty encountered in differentiating between the behaviors. Subjects who eliminated alternatives by a generalization, seemed to verify it by using the step-by-step method. More complications arose when the subjects based their step-by-step method on the generalization. If a subject generalized that white cows must give the most milk and then, using the step-by-step method, verified the generalization, one of the raters apparently checked the step-by-step method and the other indicated the generalization procedure. Although this discrepancy was discovered after only 10 tape scripts were analyzed, the differences in interpretations between the raters was not resolved. This finding suggests that the difference between the two methods of eliminating alternatives were not discernible on this activity.

Disagreement on the behavior "checked throughout the process" was apparently due to attempts to identify 3 periods during the problem-solving process where checking the results may have occurred. Raters could not agree on whether the subject checked parts of his reasoning throughout the process or just checked the answer before he gave it. Since this behavior was only observed 5 times in the "partly present" category for the subjects who gave right answers, the discrepancy passed without being identified during the analyses of tape scripts. Rater agreement of only 60 percent indicates that this behavior was not readily discernable on this problem.

"Persisted in an inappropriate method" was observed with 60 percent of the subjects who gave wrong answers to the red cows problem. The level of rater agreement was less than 47 percent. This level of agreement, although too low to indicate that it was closely related to failure in attempts to solve the problem, suggests that tape script analysis is a better means of observing this behavior than observer agreement since it was only observed 3 times and there was no agreement in the analysis of the observations. These findings suggest that use of an inappropriate method was difficult to identify on the red cows problem. The rater agreement level was 80 percent or better on the remaining behaviors.

Inspection of the data on the socks problem indicates that rater agreement ranged from 50 to 100 percent. "Emphasized key words during solution of problem" was observed as "present" in 5 cases for the subjects who gave right answers, and the raters agreed at the 100 percent level. The same behavior was observed as "partly present" in 11 cases for the subjects who gave right answers, and the raters agreed at the 91 percent level. Emphasis of key words during the solution of the socks problem was observed a total of 8 times for subjects who gave wrong answers and rater agreement only reached the 50 percent level. The disagreements in this case apparently resulted from the raters being unable to determine whether some or all of the key words were emphasized. It appeared that observations may provide a more realistic appraisal of this behavior than tape scripts since inflection of voice and short pauses cannot be duplicated on the latter.

Raters apparently failed to agree on the methods used to eliminate alternatives. It is interesting to note that rater agreement was 75 percent for "elimination by generalization" with subjects who gave wrong answers and only 50 percent with subjects who gave right answers. Apparently subjects employed the techniques of eliminating alternatives by generalization and through the step-by-step method, but differences between these techniques were not easily discernible.

It seemed obvious that subjects were required to make the essential interpretations and relationships to reach a correct solution. The analyses of tape scripts indicated that over 68 percent of the subjects verbalized these procedures. Agreement between raters averaged 82 percent. Rater agreement exceeded 97 percent on this behavior when the percentages of subjects who made none of the essential interpretations were considered. It seemed that the analyses of tape scripts as a means of determining the number of subjects, and the degree to which they stated the essential interpretations and relations, may represent an area in which good and poor problem solvers differ in their methods of problem-solving.

Problem 11 (railroad) was the third problem selected for tape script analysis. This problem seemed to confront each of the students with a difficult challenge. Mean times for correct answers were greater than 100 seconds for each grade level and only 14.5 percent of subjects gave right answers to this problem. Since the time for correct answers to the railroad problems were greater than for any of the other problems and it was the most difficult problem, it seemed that subjects would employ more techniques to arrive at an answer.

Percentages of agreement beyond the 80 percent level were found on "emphasized key words in the statement of the problem," "eliminated alternatives step by step," "stated necessary relationships," "made false interpretations," and "persisted in an inappropriate method" for subjects who gave wrong answers.

Inconsistencies in the raters' identification of the behaviors employed by the subjects seemed to be a result of the raters' disagreement as to whether behaviors occurred while the subject was inspecting the problem or while the subject was in the process of solving it. The degree to which the behavior was employed by the subject seemed to be another point of disagreement for the raters.

Inspection of the results of the analysis of tape scripts indicated that subjects who gave wrong answers were not as thorough in their attack on the problem. The students who gave right answers reviewed the problem more, made fewer mistaken interpretations, stated all the necessary relationships and stated all the essential interpretations. Subjects who gave wrong answers failed to follow their procedures of problem-solving through to completion. They only stated some of the key words, some of the relationships, some of the interpretations,and never really appeared to get all the facts together for one problem. This interpretation is based on the assumption that each of the subjects could see the necessary relationships and essential interpretations if they thoroughly explored the information given in the problem. As in the analyses of observations, there seemed to be a slightly different pattern or approach to each of these problems. It seemed that the processes used in solution were similar, but the means or points at which the subjects chose to begin solving the problem were determined by the subject.

General Summary

The following statements summarize the findings of this study. When the term "significant differences" appears below it should be interpreted as beyond the 5 percent level.

Grade Level Performances

Only 3 of 12 problems were found to elicit significant grade level differences in numbers of correct answers. In each of these problems the seniors gave more correct answers.

Sex Differences

Significant differences were found between the performances of boys and girls in 4 problems. The boys gave more correct answers to 2 of the 12 problems and the girls gave more right answers to 2 family-relationship problems.

Confidence

Analyses of the responses to the confidence question indicated differences between the sexes. In one-half of the problems more boys than girls indicated they were sure their answers were correct. Girls indicated that they were doubtful of their answers more frequently than boys on 4 problems. More boys than girls indicated they were sure their answers were *in*correct on only one problem.

Relationships Between Objectively-Scored Tests and Experimental Problems

Total scores on each of 4 objectively-scored tests were significantly related to at least one of the experimental problems. Five significant coefficients appeared between the experimental problems and each of 2 objectively scored tests. Scores on 2 other tests were each significantly related to 4 of the experimental problems.

Observed Behavior

1. More subjects who employed the if-then approach gave correct answers to 5 of the 12 problems.
2. In 5 of the problems more subjects who enjoyed the challenge gave correct answers.
3. Observers indicated that subjects who checked their reasoning before giving a final answer gave more correct answers to 3 problems.
4. When the subjects used their hands during the problem-solving they gave more correct answers to 2 problems.
5. The subjects who employed the procedure of estimating an answer as a basis for solving a problem gave more correct answers to one problem.
6. It was found that more subjects who guessed gave wrong answers to 2 problems.
7. Subjects who persisted in an inappropriate method gave more wrong answers to one problem.
8. Observers indicated that subjects who brought in outside distractors into the problem situation gave more wrong answers to one problem.
9. In only one problem the subjects who rejected their first attempt to reach a solution gave more wrong answers.

Analyses of Tape-Scripts

Analyses of tape-scripts corroborated the results of the analyses of observed behavior. It was found that subjects who stated more key words, made more interpretations, saw more relationships, and used fewer false interpretations gave more right answers.

Summary and Implications

Two hundred superior high school students, 25 boys and 25 girls from each of ninth, tenth, eleventh and twelfth grades in 50 schools were subjects of this investigation. All of the subjects were participants in the Research and Guidance Laboratory for Superior Students at the University of Wisconsin. Each was confronted with 12 problems which they were to solve while thinking out loud. The behavior exhibited during the activity was observed and recorded by the administrator of the exercises. Seventy of the subjects were observed by an observer in addition to the administrator. Tape recordings were made while the subjects were solving the problems, and independent analyses of them were made by two researchers. Detailed reports of the findings have been reported in the body of the report. The general conclusions with their implications are noted below.

The differences between percentages of subjects at different grade levels who gave right answers were significant in only 3 problems. The senior students gave more correct answers than freshmen in 2 problems and gave more correct answers than the ninth, tenth and eleventh graders in one problem. In the 9 remaining problems, no significant differences appeared. It appeared that high school seniors were not superior to underclassmen in solving the problems used in this investigation. High school personnel might investigate these results in relation to their stated objectives of teaching students how to solve problems. Since many of the errors made by the students seemed to be of the kind that could be corrected by instruction, an increased effort to provide instruction in "thinking on their feet" might enable students to improve their skills.

Differences between the percentages of correct answers by boys and girls in 4 of the 12 problems were found. Two problems were answered correctly more often by girls than boys. The boys gave more correct answers than girls to 2 problems. Educators should be aware that sex differences may occur in problems which seem to provide equal opportunities for solution by boys and girls. The material in the problem and its structure should be examined and perhaps tested for sex differences before assuming that boys and girls are equally capable of giving correct answers.

Differences between boys and girls appeared when the subjects were asked to express their confidence in the correctness of their answers. The male subjects stated, more than the girls, that they were certain of the correctness of their answers in half of the problems. The girls indicated more frequently that their answers to 5 problems might be wrong. Few students suggested that they were sure their answers were not correct. Counselors and teachers could use information relevant to sex differences in confidence in working with individuals. It seemed that superior students may be confronted with opportunities to express their confidence in different problem situations. The awareness of possible differences between boys' and girls' behavior might aid counselors and teachers in helping superior students to understand themselves.

The degree of relationship between total scores on the experimental exercises and scores on objectively-scored tests was low. This may be due to the fact that the reliability of total scores from the 12 exercises was low. It appeared that right and wrong answers to the exercises alone provided little additional information that could not be obtained from total scores on the objectively-scored tests. High school personnel might profit from experimental results which indicated that simple evaluation of the correctness of answers to problems did not discriminate among superior high school students. The suggestion seems to be that more than the *products* of reasoning should be evaluated when counseling and teaching superior students.

Requiring students to think out loud during problem-solving seemed to be one means of evaluating the processes of reasoning. When 2 observers recorded the behavior exhibited by the same subjects, the agreement was generally high, but the extent of disagreement suggested that the behavior depended partially upon the nature of the problems. When differences between the procedures of subjects who gave right and wrong answers were analyzed, it was found that subjects who gave correct answers were more likely to use the *if-then* approach and to *enjoy the challenge* of the exercise. In 5 problems these differences were significant. *Estimating an answer* as a guide to solving the problem was employed by more subjects who gave right answers on one problem. Subjects who *used their hands* during the attempted solution process were more likely to answer one of the most difficult problems correctly. These findings cannot be generalized but the evidence suggested that procedures employed during effective problem-solving could be observed. Information obtained during observation might be useful in helping individuals to note their weaknesses and to correct them.

Subjects who gave wrong answers were more likely to *guess*, to *bring irrelevant information into the problem situations*, and to *persist in an inappropriate method*. In each exercise observers reported that more subjects who made such errors gave wrong answers. Correction of minor errors in problem-solving might aid superior students in developing their apparent potential for good reasoning.

Raters who examined the protocols typed from tape recordings identified practices similar to those noted by observers. Subjects who gave right answers stated *all of the key words*, *all of the essential interpretations*, and *all of the necessary relationships*. Wrong answers were more likely to be given by subjects who *made false interpretations* and stated only some of the *necessary relationships and interpretations*.

Differences in the degree to which procedures were employed suggested that the less skilled problem-solvers were not as thorough as those who were effective in their attempts to find solutions. Perhaps they had not learned to assess _the value of each part of the problem_ in reaching a solution. It seemed that this type of information could be used effectively by teachers and counselors in helping students to appraise themselves.

It is frequently suggested that the function of the counselor is to help all young people to help themselves to solve the problems that arise inevitably in the process of coming of age in America. This exploratory study of problem-solving suggests that further work in this area may produce a technique to help students in discovery of their strengths and weaknesses in problem-solving. The findings suggest that analyses of counselees' performances and interpretation of the results may provide a valuable addition to the counselor's techniques.

CHAPTER IV

PROBLEMS IN ACTUARIAL AND CLINICAL PREDICTION OF PERFORMANCES OF SUPERIOR STUDENTS

The third study was concerned with problems in use of actuarial and clinical methods in forecasting of post-high school performances of superior secondary school students.

Review of the literature had revealed that 3 important reports on this problem had been published. Meehl's comprehensive discussion (1954) of the problems involved in prediction is particularly significant since it raises many issues in the processes employed. Meehl and Rosen's 1955 report on antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores raised the issue of the extent to which investigators must recognize the probable occurrence of events over which they have no control. The report by Chamberlin and others in 1942 on the success of subjects of the Eight Year Study (in which this writer participated) contained an excellent presentation of criteria of success in college and the means by which data about them could be obtained and appraised.

Hundreds of studies of prediction of college students' first-year grade-point averages by means of test scores and high school marks have been conducted from the time of Dearborn's pioneer study in 1909 down to the current efforts of the National Merit Scholarship group in 1959. Coefficients have clustered in the .40 to .50 range. The results are not particularly useful in counseling the individual student because the studies cover too short a criterion period, and the use of grade-point averages as the criteria omit important areas of evidence about achievement in college. Rothney's (1958) attempt to predict success over the 4-year period of college attendance by regression equations based on high school grade records, test scores, measures of socio-economic status and, separately by clinical judgments made on a 5-point scale at the time of high school graduation, indicates that the statistical and clinical predictions were comparable. Addition of clinical and statistical data did not improve the forecasts significantly. Similar findings were obtained when statistical and clinical predictions of progress in employment were made at the time of high school graduation. The criteria used were reports of 100 of the subjects 5 years after graduation and verification by employers of the subjects' statements.

Polansky (1941), Estes (1938), Allport and Fredericksen (1941) used small numbers of subjects in clinical predictions of performances or ratings and claimed that their methods were effective. Neither of these authors made comparisons of the efficiency of clinical forecasts with that obtained by actuarial methods. Sarbin (1941) attempted such a comparison in predicting grade-points of college students and found no important differences between the 2 methods. Conrad and Satter found, in 1945, slightly better actuarial than clinical predictions of the success of 3500 trainees in a Navy electrician's school. The results of attempts to forecast success of aviation cadets as aircrew members in World War II (in which this writer participated) indicated that the clinicians did not predict performances as well as the statisticians. Kelly and Fiske (1950) suggested that both statistical and clinical approaches worked equally well in forecasting the performances of trainees in clinical psychology but there was some contamination due to the fact that clinicians were familiar with their subjects' test scores. Bloom and Brundage (1947) showed that when interviewers had test scores of their subjects (naval enlisted men) available during interviews their forecasts of success in training were lower than those made by investigators who used only test scores in making predictions.

PROCEDURES AND RESULTS

In June, 1960, the first group of subjects who had been sent to the Laboratory graduated from high school. They constituted a small sample that was later found to be unrepresentative of the students who were to be sent in later years. Accordingly, the graduates of 1961 were used as the trial groups for this study and the 1962 graduates who went to the University of Wisconsin were used as a sample for a comparative study. Unfortunately the period covered is short but, as these subjects are followed through many years of post-high school training and occupations it will be possible to determine effectiveness of various predictive procedures over much longer periods of time.

At the time each subject graduated from high school and indicated the college that he or she planned to attend, a member of the Laboratory staff studied his complete record and made a clinical prediction about his success in post-high school training. Samples of the predictions for 2 graduates, edited to prevent identification, follow.

Prediction of Performances of Albert in Post-High School Training

Name Albert
Place of training Uni. of Wis.
Counselor Jones
City ______
Date ______

He/she will:

- ___ A. be eligible for major awards based upon scholastic performances.
- ___ B. be eligible for minor awards based upon scholastic performances.
- X C. complete training. Will not be eligible for awards nor on probation.
- ___ D. take longer than usual time to graduate though never on probation.
- ___ E. drop out of training for reasons other than scholastic difficulty.
- ___ F. be a marginal student. Will be on probation at some time.
- ___ G. dropped because of low grades.

Academic performance. Has been inconsistent. Overall average is approximately C+. Best performance has been in mathematics and physical sciences.

Test performances. Test performances are very high in all areas. Never below 95th percentile with most scores placing him in the 99th. Raw score of 73 on Concept Mastery Test. National Merit semi-finalist.

Strength of drive. Al achieves well in subjects which interest him. Seems to have a tendency to give routine effort to the "necessary" subjects in his program which exert no great appeal for him. He is going to college, however, because he wants to. Has a clear goal in sight. Was intent on spending time with head of mathematics department at a university and with personnel in Numerical Analysis Laboratory. Estimate that he has drive but not of real intensity.

Realism in self-appraisal. Seems to have selected a goal congruent with his interests and opportunities.

Flexibility, adjustability, modifiability. This is a question mark with Al. Has been described by faculty and self as one who loves to take the least popular stand on an issue purely for the sake of being different. Sticks tenaciously to his position in such cases even without adequate data to support it. Primarily, however, the question is: Will he apply himself to English, history, etc. or will his interests dictate the quality of his work?

Handicaps. Rather pronounced case of acne. Is very slim and almost frail. Tires easily but no organic trouble is indicated.

Home circumstances. Three younger siblings. Father owns small business. Al will be able to finance his education, but is considering part-time work. Scholarship aid would probably enable him to give full-time to study.

Unique factors. Has associated primarily with students older than himself. Has spent much of his free time talking with and observing personnel at a Numerical Analysis Laboratory.

Prediction of Performances of Mary in Post-High School Training

Name Mary
Place of training Uni. of Wis.
Counselor Smith
City
Date

He/she will:

X A. be eligible for major awards based upon scholastic performances.
___ B. be eligible for minor awards based upon scholastic performances.
___ C. complete training. Will not be eligible for awards nor on probation.
___ D. take longer than usual time to graduate though never on probation.
___ E. drop out of training for reasons other than scholastic difficulty.
___ F. be a marginal student. Will be on probation at some time.
___ G. dropped because of low grades.

Academic performance. Her performance in school has been consistently at the highest level. From all indications, she learns quickly and has established good study habits.

Test performances. Her test scores have been above the 95th percentile. The raw score on the Concept Mastery Test was 65 which is near the average for Laboratory students.

Strength of drive. Mary's drive appears to be stimulated by a mature outlook and a desire to explore academic areas for knowledge itself rather than for specific information required for a course.

Realism in self-appraisal. She seems to appraise herself realistically. At this stage in her development she questions that she can excel in academics. It appears that these doubts will serve as motivation rather than a hindrance to success.

Flexibility, adjustability, modifiability. Mary's varied interests and activities indicate that she is a flexible girl, who adjusts to most situations. Her behavior, as observed at the Laboratory, suggests that she can adjust to most situations with ease.

Handicaps. She has displayed no handicaps at the Laboratory which might be obstacles to her success in college.

Home circumstances. Mary's home environment appears to be conducive to achievement and adjustability. Father is in a profession. One older sister in college. Needs financial help.

Unique factors. She has unusual poise. She has displayed social and intellectual maturity which stands out even among the superior students at the Laboratory.

Actuarial Predictions for 1961 High School Graduates.

During their visits the students had taken the following tests in the grades indicated. (It should be noted that in the first years of the Laboratory several tests were tried out. Experimentation has indicated that, for several reasons, many were not useful and their use has been discontinued. Others are being tried.)

Ninth grade	School and College Ability Test Differential Aptitude Tests of Abstract reasoning Language usage Spelling
Tenth grade	Differential Aptitude Test of Verbal reasoning Sequential Tests of Educational Progress in Reading
Eleventh grade	Sequential Tests of Educational Progress in Social Sciences Sequential Tests of Educational Progress in Natural Sciences
Twelfth grade	Sequential Tests of Educational Progress in Mathematics Terman Concept Mastery Tests

In addition to the test scores 2 measures of academic performances were available. A complete record of each student's marks in grades 9 through 12 was obtained and a high school grade-point average was computed. Verified records of the first semester marks of all graduates were obtained from those students who attended the University of Wisconsin after high school graduation, and students' reports of performances who entered other colleges throughout the country were procured. These first semester grades were the criteria in the prediction studies. (Cumulative records of grade-point averages as the students proceed through college are being maintained at the Laboratory and it will be possible to use them in prediction studies over longer periods of time.)

Correlation coefficients indicating relationships among test scores and college and high school grade-point averages appear in Table 1. It becomes apparent that the high school grade-point averages were the best predictors of college grade-points and that the verbal test scores, particularly those from the Terman Concept Mastery Test, given in the senior year of high school, were second best. The DAT

Table 1. Coefficients of Correlation Between Test Scores and High School and College Grade-Point Averages of Eighty-nine Superior Students who Entered Colleges in September, 1961

Items	Col. grade point	H.S. grade point	SCAT verb.	DAT verb.	DAT lang.	DAT spell.	STEP read.	Terman Concept			SCAT quant.	DAT abstract
								I	II	Total		
College grade point	-	.62	.34	.20	.37	.19	.17	.40	.34	.39	.23	- .02
H. S. grade point	-	-	.30	.13	.30	.16	.00	.34	.29	.34	.20	- .03
SCAT verbal	-	-	-	.57	.58	.36	.44	.66	.54	.66	.24	.17
DAT verbal	-	-	-	-	.33	.25	.20	.52	.47	.52	.29	.23
DAT language	-	-	-	-	-	.51	.33	.48	.46	.49	.22	.18
DAT spelling	-	-	-	-	-	-	.20	.35	.35	.36	.25	.06
STEP reading	-	-	-	-	-	-	-	.40	.29	.39	.01	.18
Terman C. M. I	-	-	-	-	-	-	-	-	.69	.95	.25	- .02
Terman C. M. II	-	-	-	-	-	-	-	-	-	.86	.25	.17
Terman C. M. Total	-	-	-	-	-	-	-	-	-	-	.23	.06
SCAT quant.	-	-	-	-	-	-	-	-	-	-	-	.24
DAT abstract	-	-	-	-	-	-	-	-	-	-	-	-
Mean raw score	2.89	3.62	46.1	37.4	50.0	75.6	59.5	32.2	37.6	69.8	42.2	40.1
S.D.	0.7	0.4	6.0	5.3	13.1	16.5	4.5	14.6	8.6	21.1	3.6	4.3

abstract scores are obviously least effective of the tests for forecasting college achievement. The multiple R obtained from this battery of actuarial materials is .67. The total picture is similar to those found previously by this writer for a cross-section of high school students even though the variability for this group is less. (See Rothney, J. W. M., Guidance Practices and Results, N. Y., Harper & Bros., 1958.)

Clinical Predictions for 1961 High School Graduates

The effectiveness of counselors in predicting the performances of the same subjects is indicated in Table 2. It is apparent that they were not highly successful in their efforts. In 35 of the 89 forecasts the counselors predicted successfully for a percentage of only 39. Their predictions of minor honors for nearly half (41) of the 89 subjects was too optimistic for 17 and too pessimistic for 12. On the other hand, 9 of 22 who were predicted to fall in the average category exceeded the forecasts. In terms of a contingency coefficient, the relationship is .21 from a possible figure with this size categorization of .91.

Table 2. Relationship Between Counselors' Predictions of College Grade-Point Averages and Actual Performances of Eighty-nine Superior Students who Graduated from High School in 1961

	First Semester College Grades						
Counselors' Predictions	Drop Out	Less Credits	Probation	No awards No probation	Minor Honors	Major Honors	Total
Major honors (3.50-4.00)	-	1	-	2	3	11	17
Minor honors (3.00-3.49)	-	4	3	10	12	12	41
No awards; No probation (2.00-2.99)	-	-	2	11	8	1	22
Probation (Less than 2.00)	-	1	-	-	-	-	1
Less than usual credit load	-	-	2	1	2	-	5
Drop out for non-academic reasons	1	1	-	1	-	-	3
TOTALS	1	7	7	25	25	24	89

It is not possible to estimate accurately the reliability of the grades which were used as criteria for this study since they were earned at many different colleges and in a variety of courses but, since the multiple R with the combination of test scores and high school grades was .67, it must be at least that high. The estimated low reliability and the many circumstances that affect students' performances and counselors' judgments (See discussion in the writer's *Guidance Practices and Results*.) indicates that the counselors, all of whom had one advanced degree in guidance, were not good forecasters. Direct comparisons between the figures derived from the combination of scores and grades and the counselors' forecasts is not possible, but it appears that the latter are not as effective.

Actuarial Predictions for 1962 Graduates with the Second Group.

An attempt was made to compare forecasts by the actuarial and clinical methods for only those students who entered the University of Wisconsin so that the variability of grading practices might be reduced. The differences in departments within the University were, of course, still present, but it seemed likely that the restriction would reduce some of the sources of error.

Sixty-six students whose high school test records were complete entered the University of Wisconsin in September, 1962. The relationships between their high school tests and grades and their performances in the first semester at the University are indicated in Table 3. It will be seen that the highest coefficient was lower than for the 1961 group (.49 as compared with .62). The range of high school and college

Table 3. Coefficients of Correlation Between Test Scores and High School and First Semester College Grade-Point Averages of Sixty-six Students who Entered the University of Wisconsin in September, 1963

Items	Col. Grade Point	H.S. Grade Point	SCAT Verb.	DAT Verb.	Davis Read (Comp.)	Davis Read (Speed)	Terman C.M.	SCAT Q	DAT Abs.
College grade-point	-	.49	.14	.30	.11	.24	.36	.07	.14
H. S. grade-point	-	-	.01	.03	.12	.04	.10	.04	.17
SCAT verbal	-	-	-	.58	.71	.67	.68	.23	.18
DAT verbal	-	-	-	-	.57	.42	.55	.27	.38
Davis reading (Comp.)	-	-	-	-	-	.68	.71	.34	.20
Davis reading (Speed)	-	-	-	-	-	-	.63	.36	.19
Terman C. M.	-	-	-	-	-	-	-	.21	.17
SCAT quant.	-	-	-	-	-	-	-	-	.43
DAT abstract	-	-	-	-	-	-	-	-	-

grade-points was more restricted since the schools were beginning to select higher-performing students for the Laboratory. The multiple R for the data in Table 3 was .62.

Clinical Predictions for 1962 Graduates.

Comparison of the figures in Table 2 with those in Table 4 reveals that the counselors made similar predictions for the students at the University with which they were familiar. Thirty-five of the 66 were placed correctly for a percentage of 53. The contingency coefficient calculated from the data was .23 from a possible coefficient with these categories of .91. Again it appears that the counselors were less accurate forecasters than the actuarial data.

Table. 4. Relationship Between Counselors' Prediction of First Semester Grade-Point Averages and Actual Performances of Sixty-six Students who Entered the University of Wisconsin in September, 1962

	First Semester University Grades					
Counselors' Predictions*	Less Credits	Probation	No awards No probation	Minor Honors	Major Honors	Total
Major honors (3.50-4.00)	-	-	-	-	10	10
Minor honors (3.00-3.49)	2	1	7	16	4	30
No awards; No probation (2.00-2.99)	3	3	9	7	-	22
Probation (Less than 2.00)	-	2	-	-	-	2
Less than usual credit load	1	-	1	-	-	2
TOTALS	6	6	17	23	14	66

*None were predicted to drop out for non-academic reasons and none did so.

Predictions of Performances One and One-half Years After High School.

A study of forecasts by counselors of the 1961 graduates after one and one-half years of college suggested that their long-range forecasts were better than the short ones. Coefficients of .45 and .56 between forecasts and performances were significantly higher than those obtained for first semester grades.

Relationships Between Categorized Data from Cumulative Records and Grade-Point Averages in the First Semester of College.

The following brief report is presented to indicate first steps in an effort to study relationships between categorized descriptions of subjects derived from study of Laboratory records obtained while the students were in high school and their post-high school performances in college. Since it will require several years to examine this procedure thoroughly, this first brief report should be considered only as a description of a method which is under trial at the Laboratory.

Thorough examination of students' cumulative records indicated that counselors could agree in close to 90 percent of their categorizations when they classified subjects into 4 quarters with respect to each of the areas: sociability, practicality, goal orientation (motivation), and concern for the fine arts. The items used in the classifications are too many to report in detail here, but the following samples of the extremes in the categories will indicate the general treatment. It should be noted, however, that they are derived from 4-year records of development rather than from a particular scale or inventory response.

Top quarter in goal orientation.

Subject has clear plans (although they may change in direction), seeks much information about vocations, consistently plans advanced training, takes full load of courses and honors courses if they are available, attends special educational events, seeks out educational projects, homework is planned part of daily schedules, and he makes suggestions for improvement of learning opportunities.

Lower quarter in goal orientation.

Little interest in vocations nor advanced education, selects minimum program that school permits, takes no honor courses, does not attend special educational events nor work at extra projects, avoids homework whenever possible, makes no suggestions about improvement of educational opportunities, and may indicate ways in which school could be made easier.

Top quarter in sociability.

Student indicated many times that he was interested in social affairs with others and participated in many group activities. Expressed desire for social status and read socially oriented books and magazines.

Lower quarter in sociability.

Expressed no interest, and did not participate in social activities of the school. Indicated that he did not fit in well socially.

Top quarter in practicality.

Used out-of-school time effectively in working toward goals he had set for himself. Was highly aware of self in relation to situations and demonstrated that he could manage well in them. Usually held part-time jobs.

Lower quarter in practicality.

Did not demonstrate effectiveness in accomplishing even minor goals. Little planning, or effective use, of out-of-school time observed.

Top quarter in fine arts.

Participated much in activities that placed primary emphasis on the beautiful rather than the useful in pursuit of music, art, and literature.

Lower quarter in fine arts.

No participation or expressed interest in music, art, and literature.

Intellectual competence.

An average of standard scores (based on the Laboratory population) on 6 tests was computed for each student and the total range was divided into 4 quarters. High school grade-point averages were also distributed into 4 groups. Combining of the test scores and grade-point average, then, permitted the classification of the subjects into 4 quarters.

The procedures described very briefly above provided 5 classifications from the students' records. By assigning points of 4, 3, 2, and 1 for the standing within each of the categories, it was possible to assign a student a score that could range from 5 for a person in the lowest quarter of each category (only one student was so classified) to 20 for one who was classified in the top quarter. This score might be named temporarily a measure of *general development*. It was recognized, of course, that there would be overlap in the areas represented and that other objections might be raised to the procedure but, as pointed up previously, these were first exploratory steps in the development and testing of hypotheses.

The correlation between high school general development measures and first semester college grade-point averages of 100 Laboratory subjects who graduated from high school in 1961 was .38. For just those subjects who entered the University of Wisconsin the coefficient was .42. For similar groups who graduated in 1962 the coefficients were .43 and .58. The results of this exploratory study offers enough promise to warrant further investigation.

The procedures and categories used in classifying the Laboratory subjects while they were in high school have been attempted with data obtained from them as they proceeded into college. Again classification seemed to be possible with high agreement among those who attempted to classify the subjects. When enough follow-up data becomes available it seems that it will be possible to get criterion measures of college performance that will go beyond the usual grade points criteria that are commonly employed. It may then be possible to determine whether clinical evidence (clinical because classification requires interpretation that goes beyond mere counting) will be more or less effective in forecasting post-high school achievements than actuarial data.

Summary

There seems to be little doubt that combinations of actuarial data obtained from tests given in high school and grade-points earned there were better predictors of early college performances than the forecasts made by counselors who had studied the cumulative records (including tests) of their counselees. Similar findings have been obtained by other investigators. Reasons for such results have been discussed elsewhere by the writer and are not repeated here.

A new method for attacking the problem of prediction of post-high school performances of superior students on a broader base, and with criteria which go beyond simple college grade-points as measures of post-high school performance, has been developed. Preliminary trials indicate that the method warrants further study.